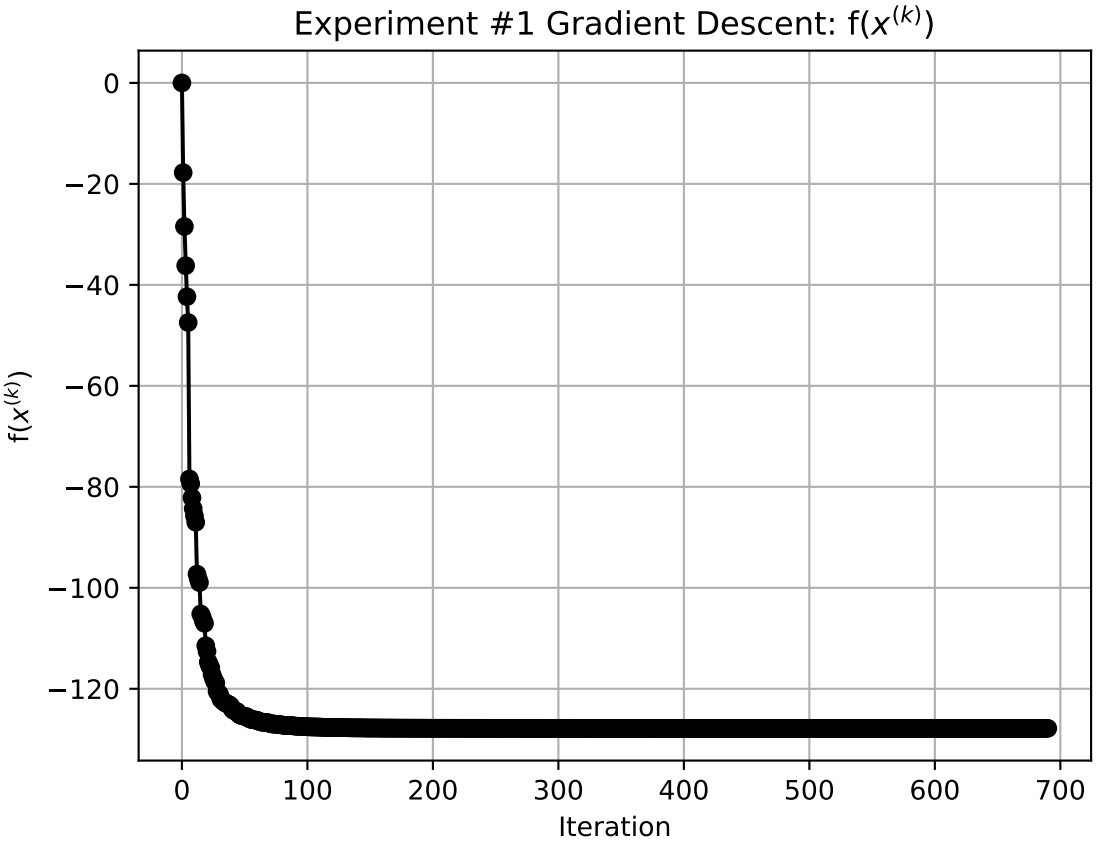
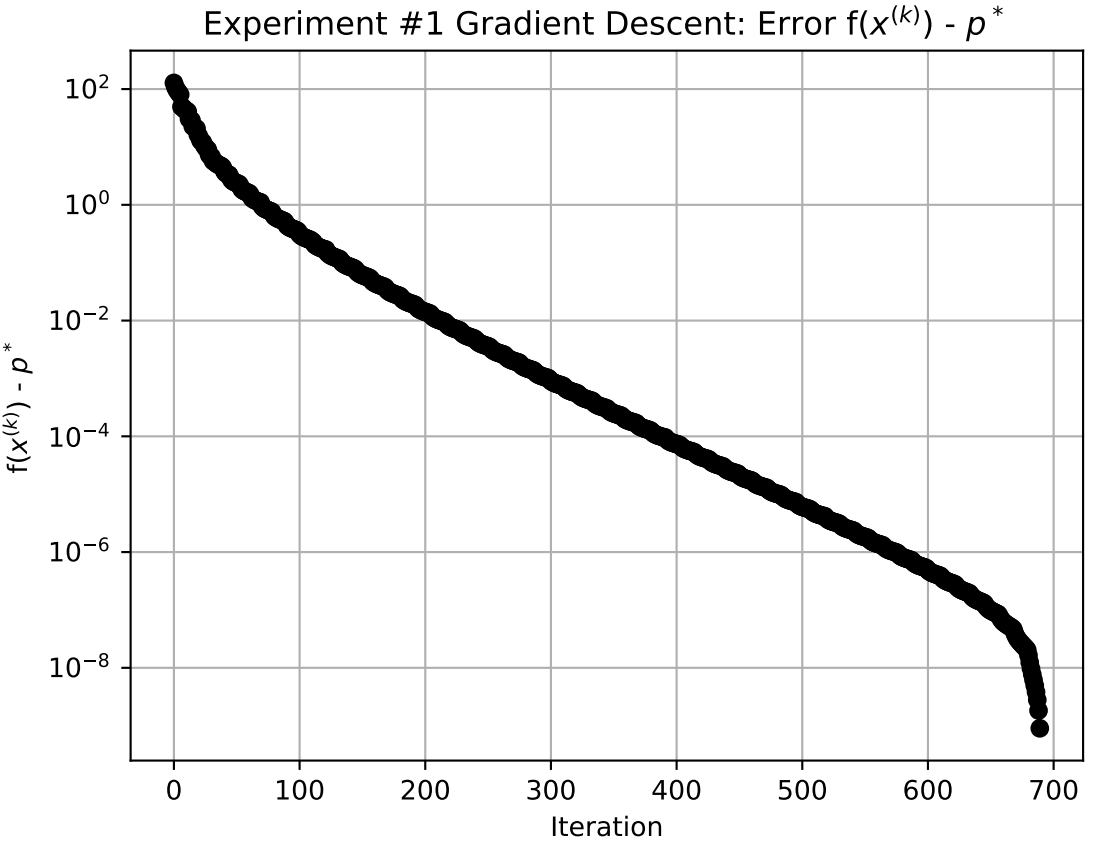
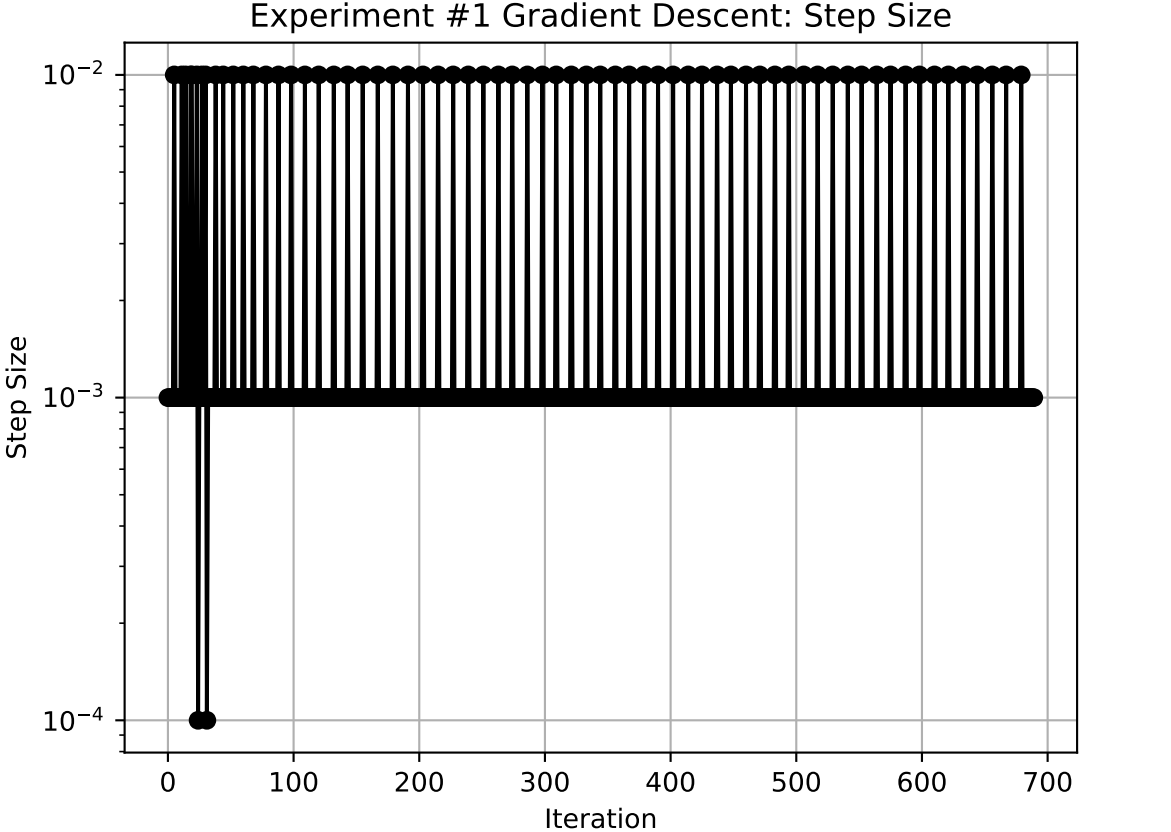
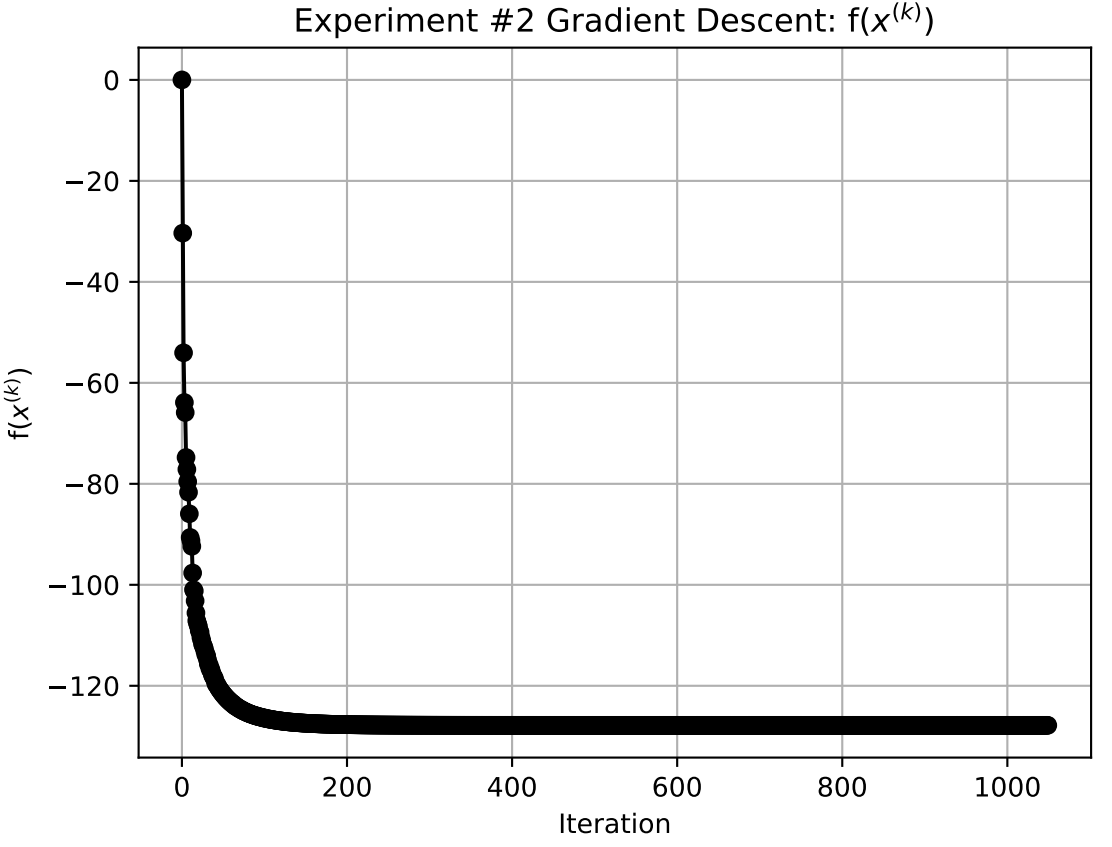
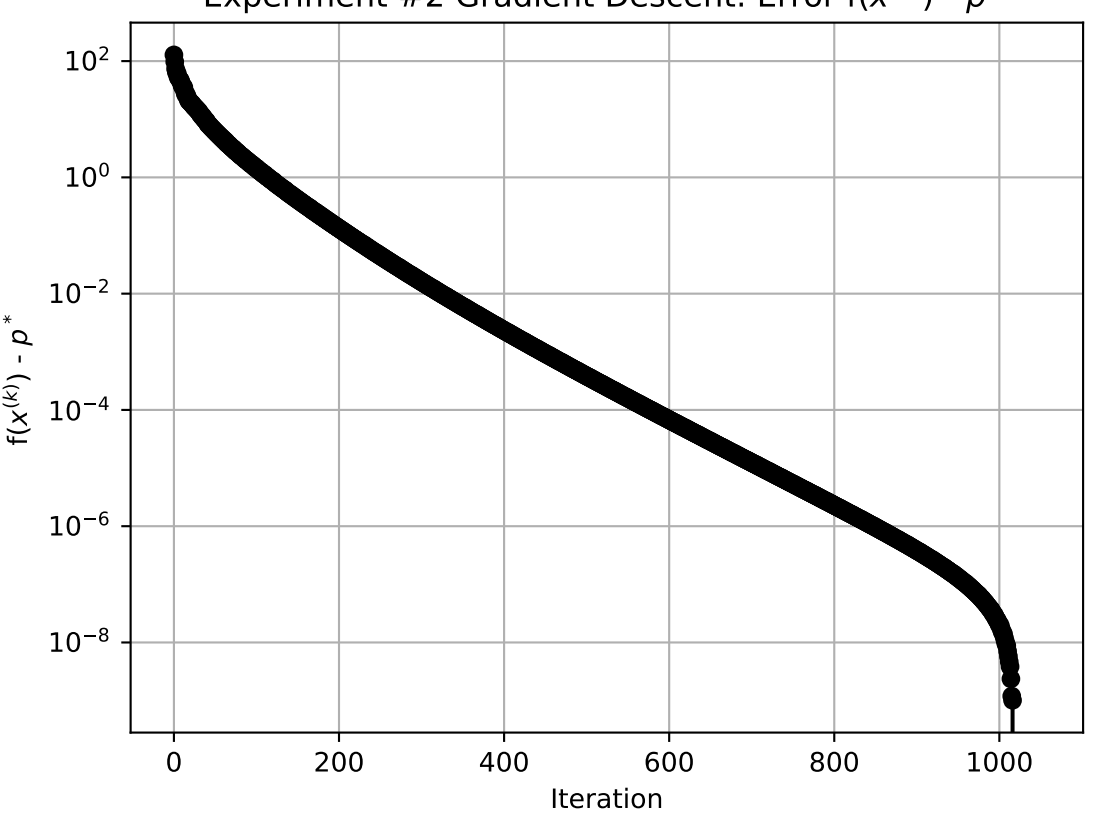
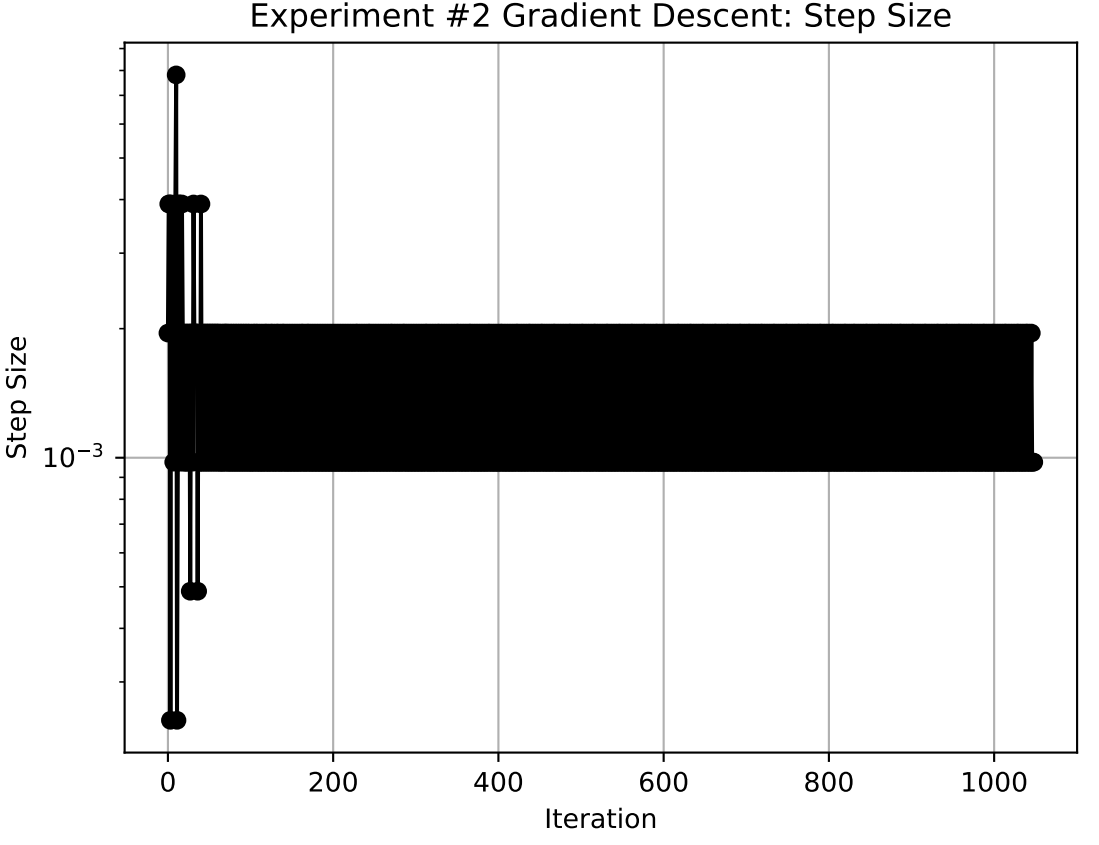

Experiment #1 Gradient Descent: f($x^{(k)}$)
0
−20
−40
−60
−80
−100
−120
f($x^{(k)}$)
0
100
200
300
400
500
600
700
Iteration

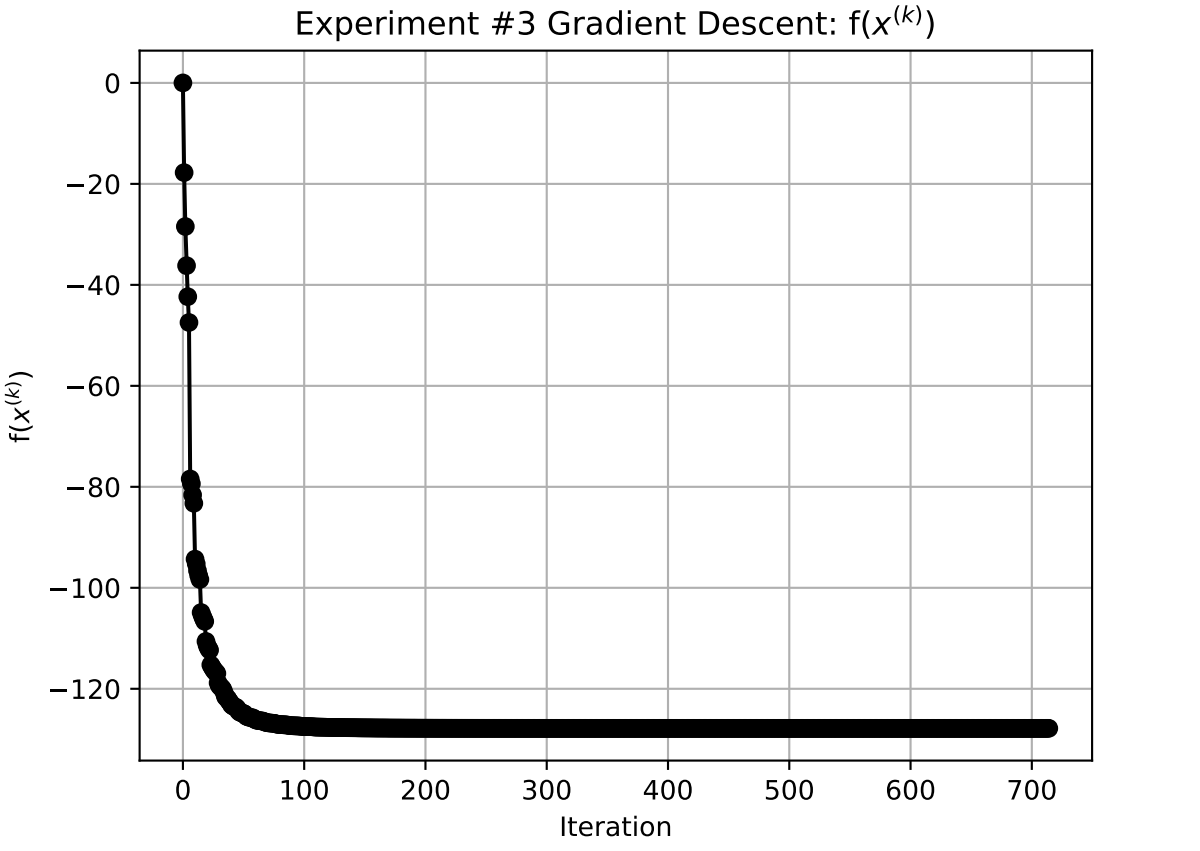

Experiment #1 Gradient Descent: Error f($x^{(k)}$) - p^*
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
f($x^{(k)}$) - p^*
0
100
200
300
400
500
600
700
Iteration

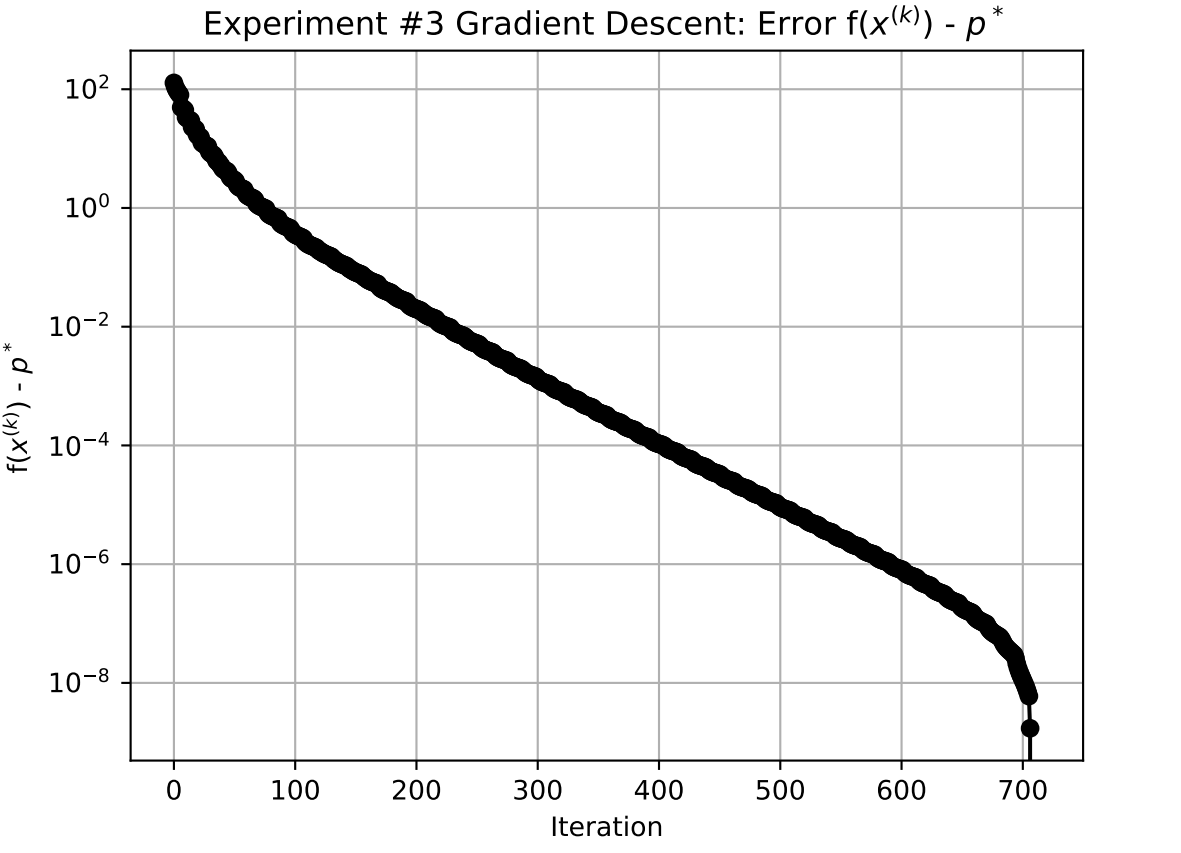

Experiment #1 Gradient Descent: Step Size
Step Size
10^{-2}
10^{-3}
10^{-4}
0
100
200
300
400
500
600
700
Iteration

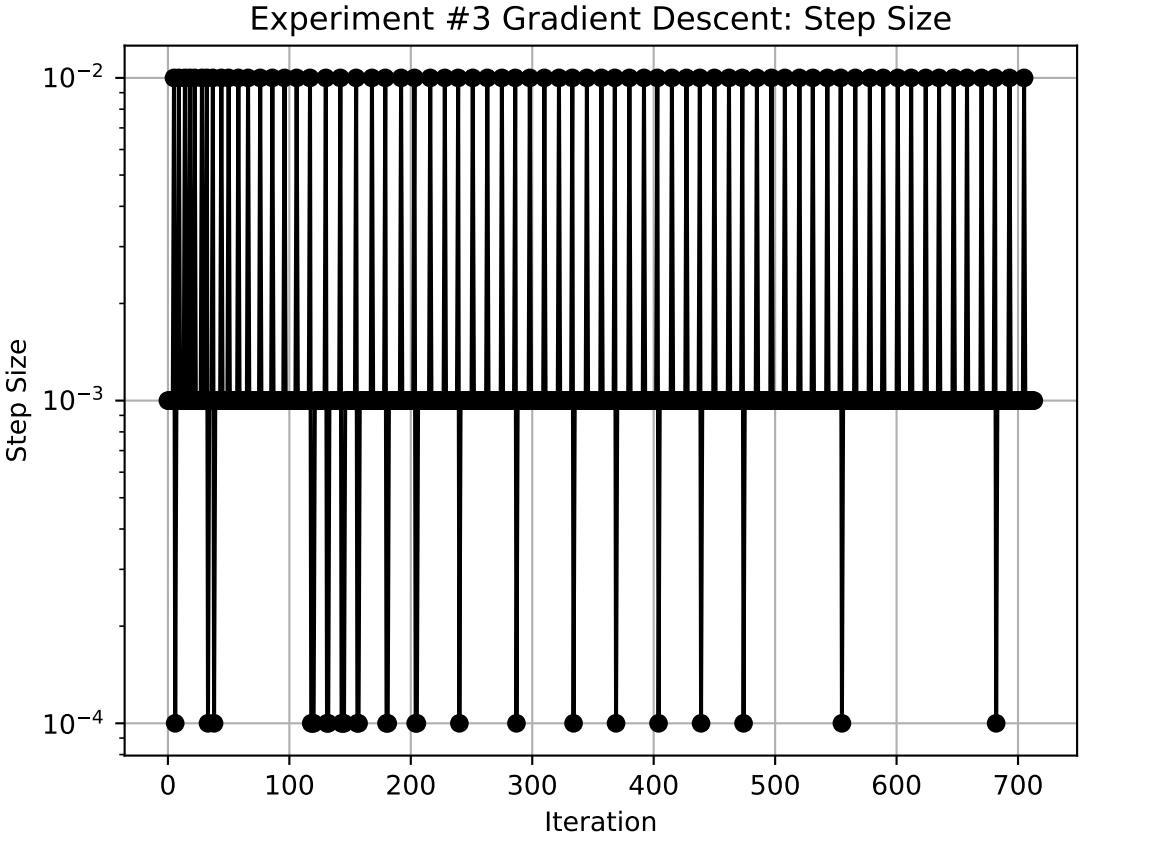

Experiment #2 Gradient Descent: f($x^{(k)}$)
0
−20
−40
−60
−80
−100
−120
f($x^{(k)}$)
0
200
400
600
800
1000
Iteration

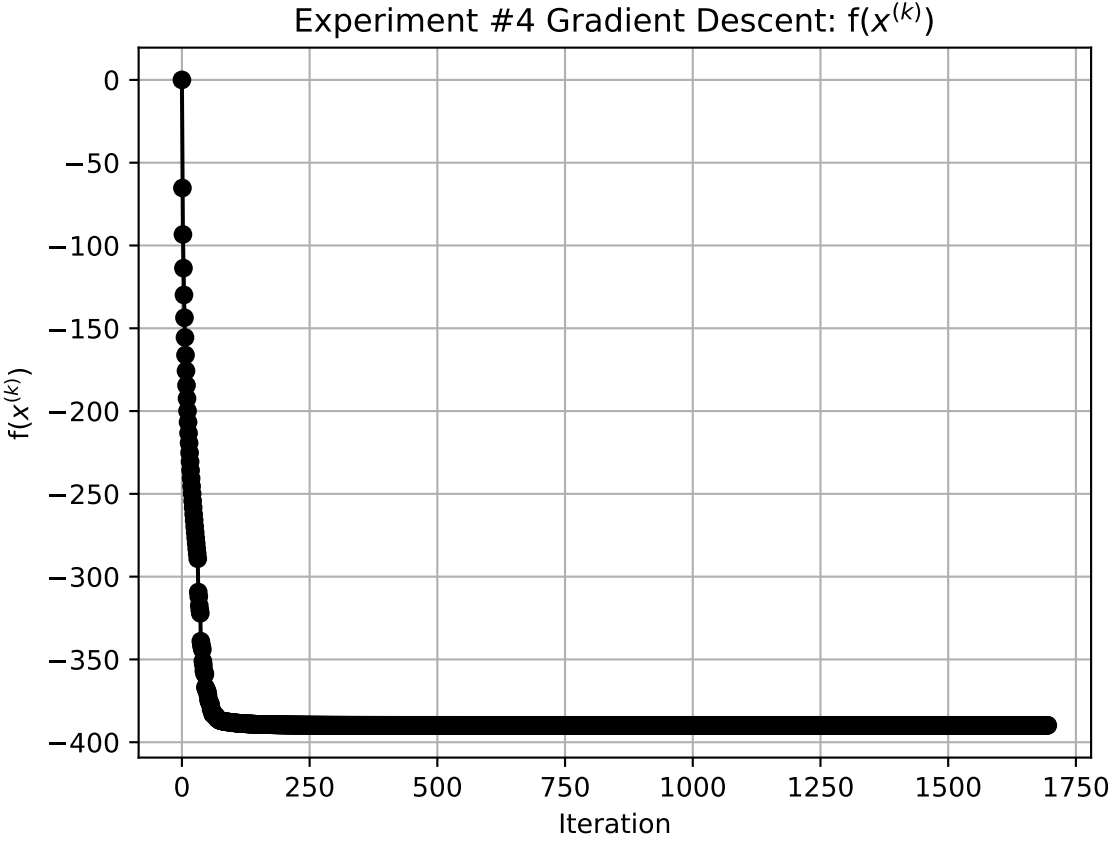

Experiment #2 Gradient Descent: Error f($x^{(k)}$) - p^*
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
f($x^{(k)}$) - p^*
0
200
400
600
800
1000
Iteration

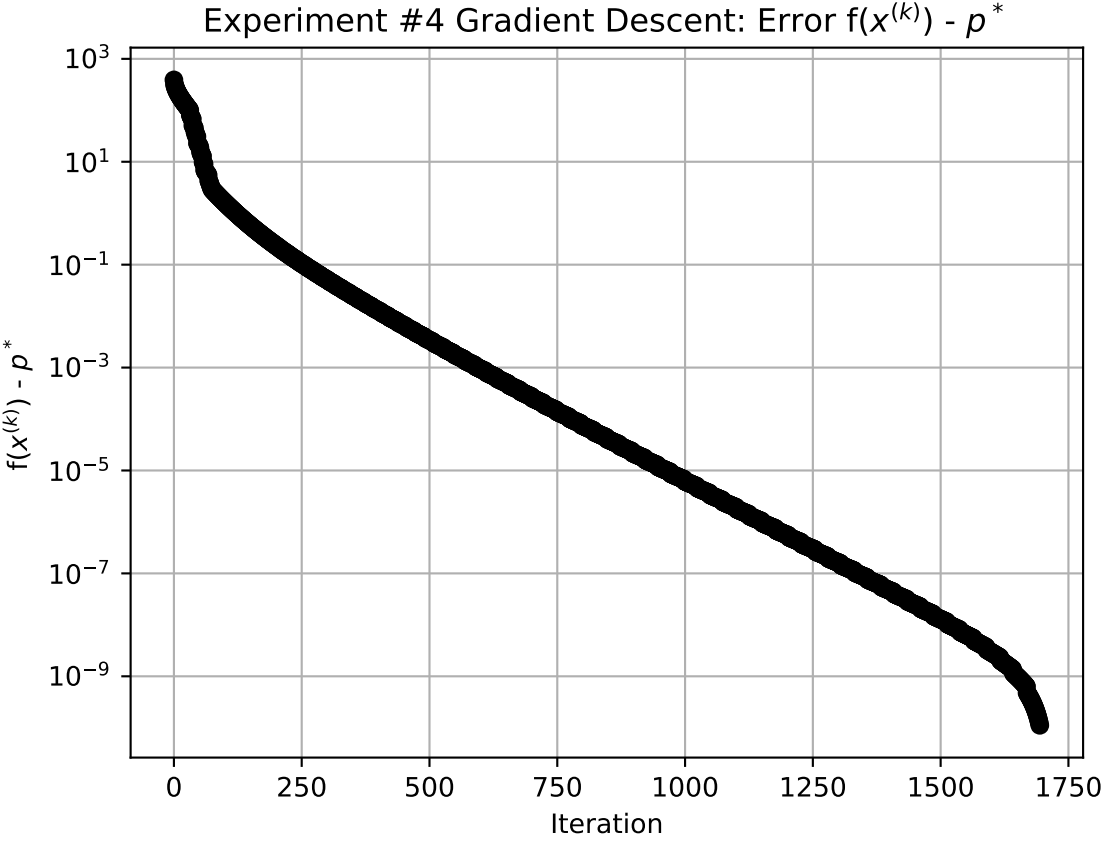

Experiment #2 Gradient Descent: Step Size
Step Size
10^{-3}
0
200
400
600
800
1000
Iteration

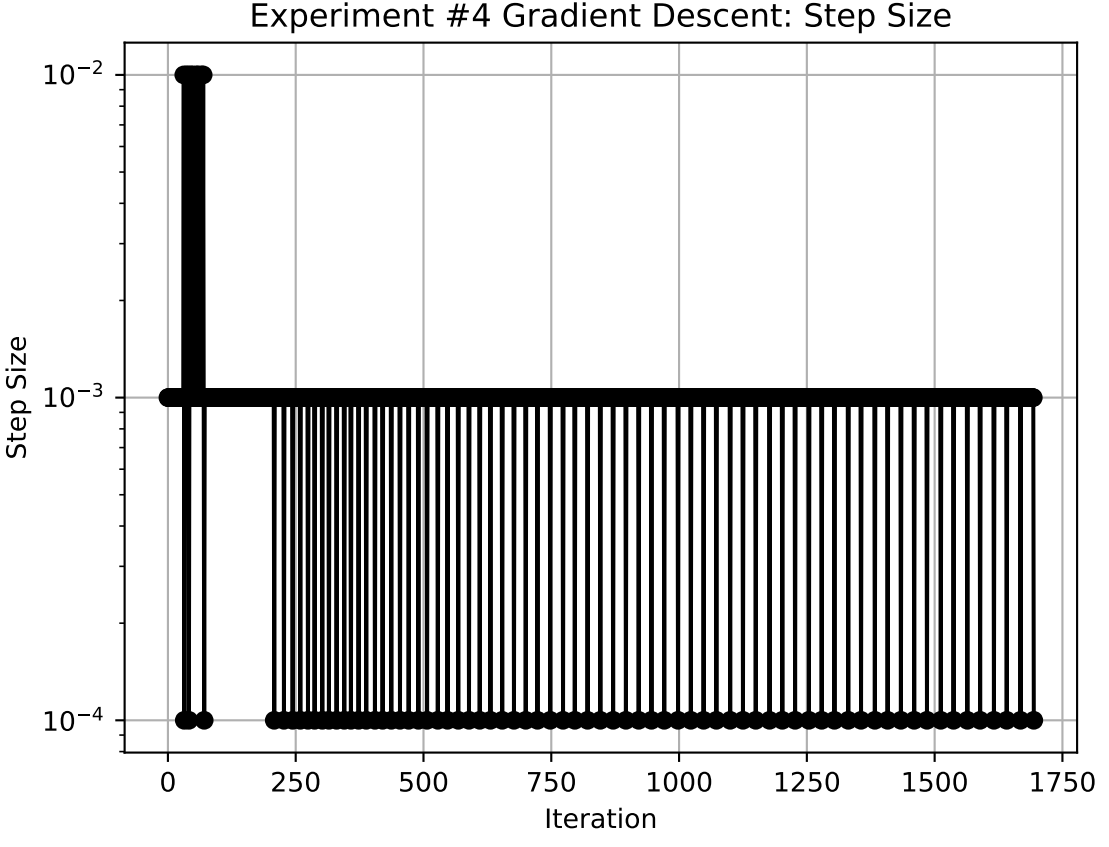

Experiment #3 Gradient Descent: f($x^{(k)}$)
0
−20
−40
−60
−80
−100
−120
f($x^{(k)}$)
0
100
200
300
400
500
600
700
Iteration

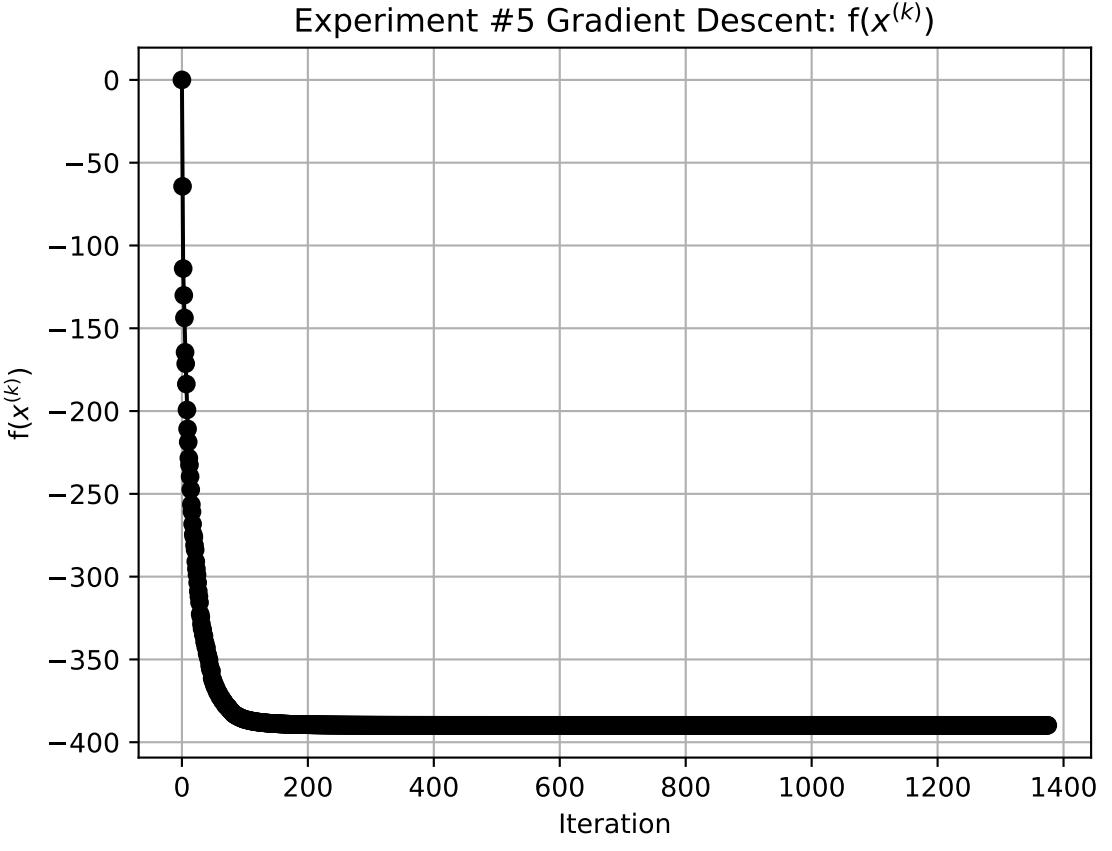

Experiment #3 Gradient Descent: Error f($x^{(k)}$) - p^*
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
f($x^{(k)}$) - p^*
0
100
200
300
400
500
600
700
Iteration

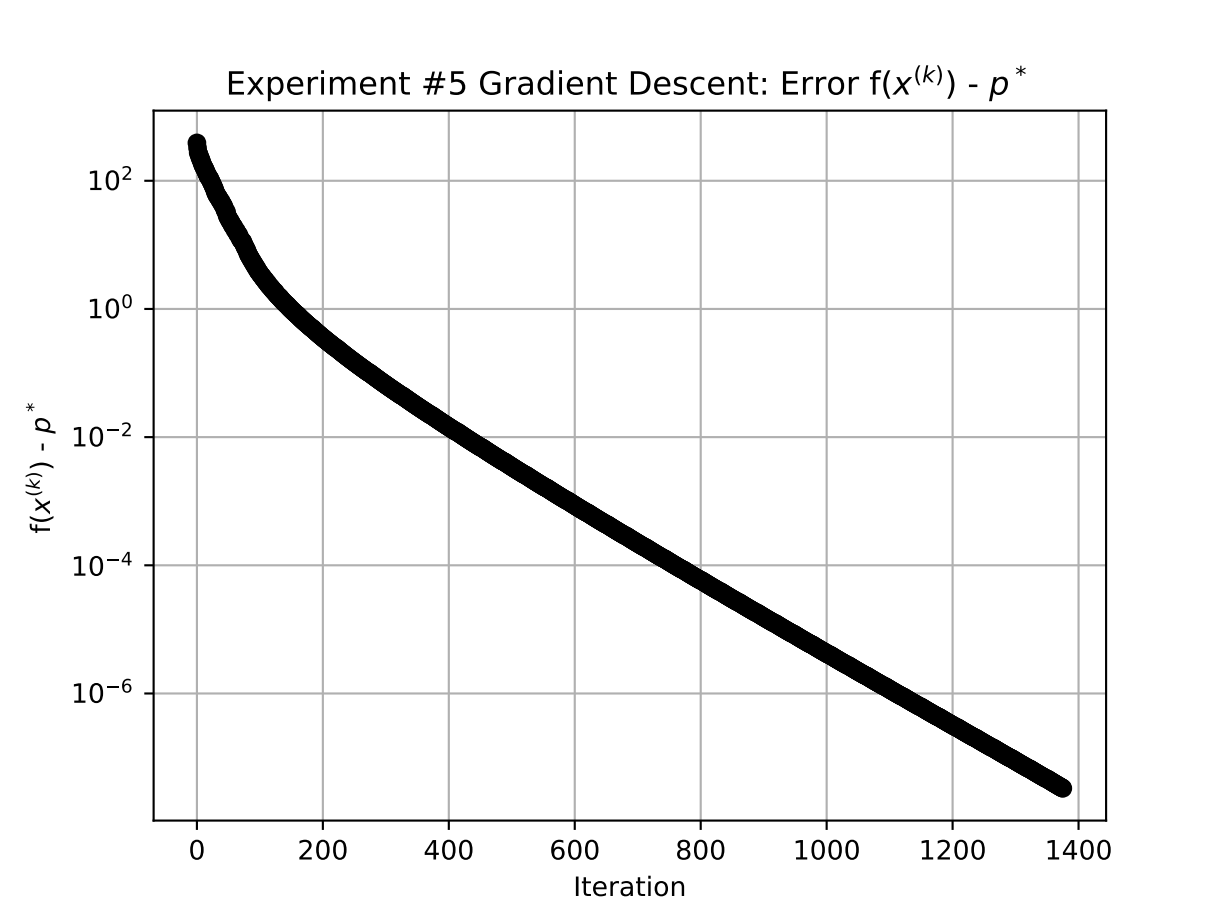

Experiment #3 Gradient Descent: Step Size
Step Size
10^{-2}
10^{-3}
10^{-4}
0
100
200
300
400
500
600
700
Iteration

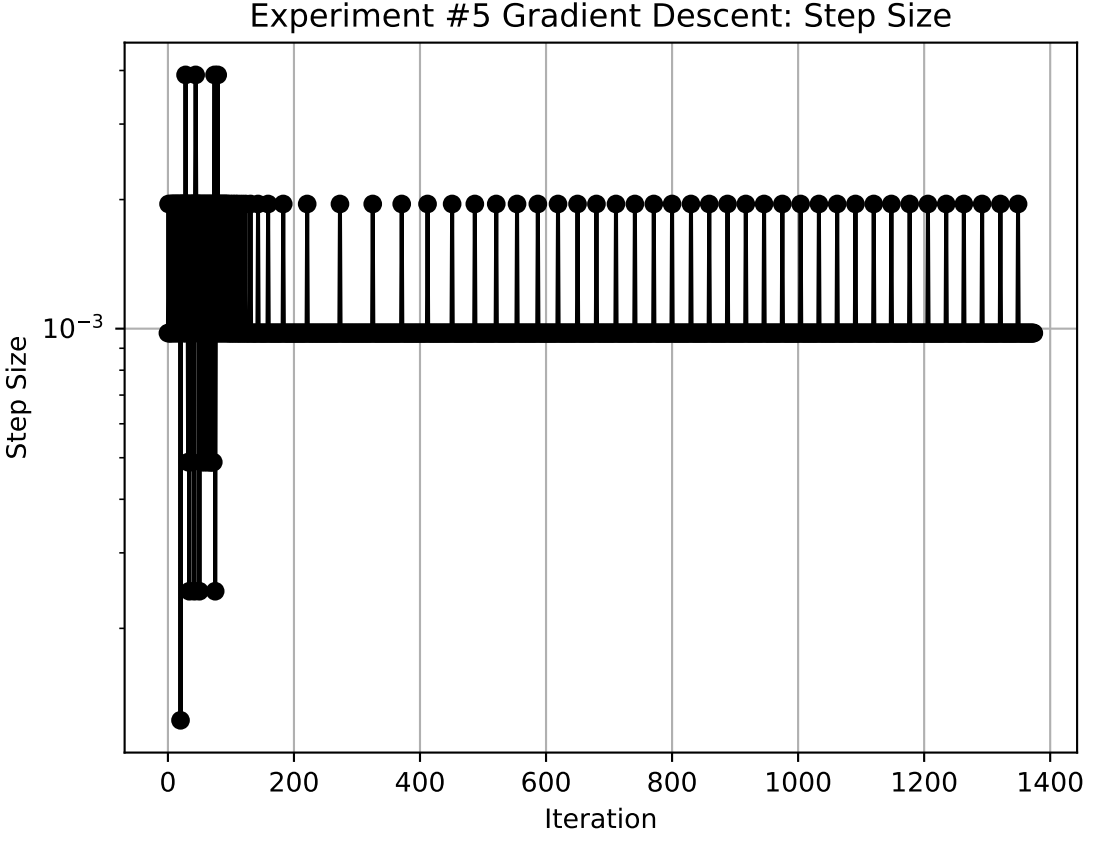

Experiment #4 Gradient Descent: f($x^{(k)}$)
0
−50
−100
−150
−200
−250
−300
−350
−400
f($x^{(k)}$)
0
250
500
750
1000
1250
1500
1750
Iteration

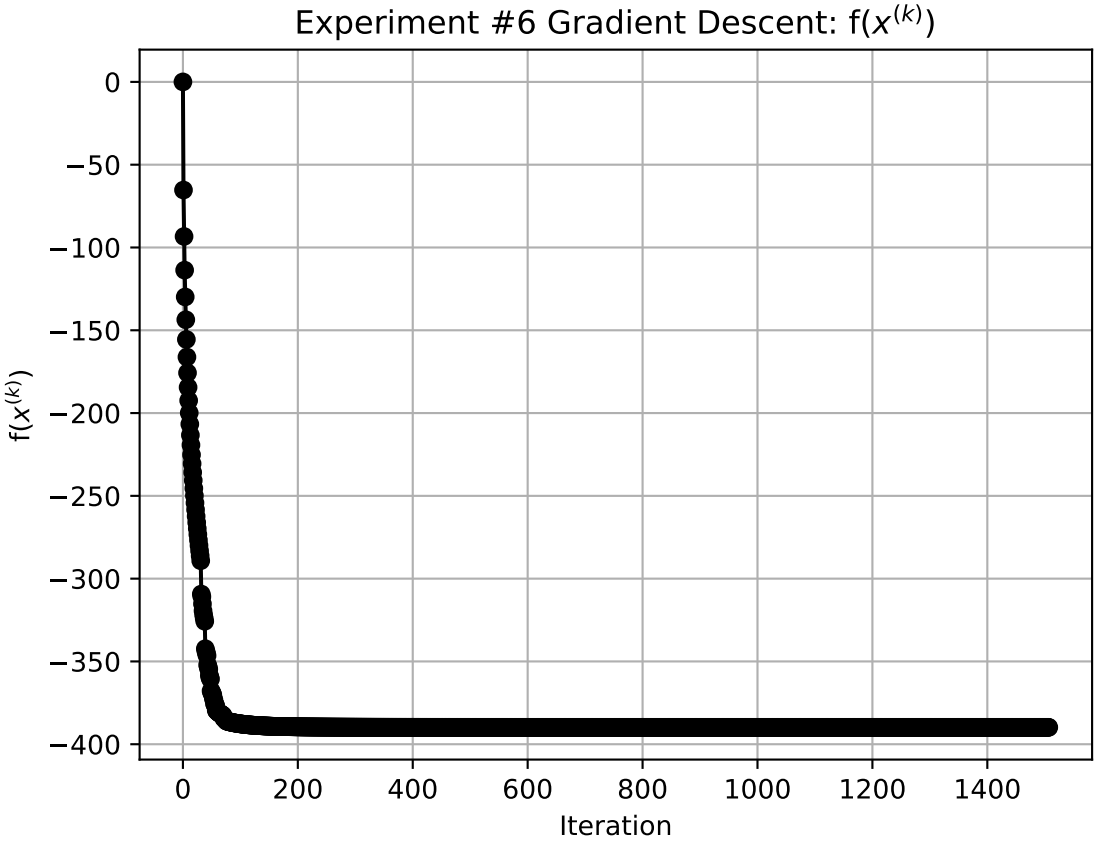

Experiment #4 Gradient Descent: Error f($x^{(k)}$) - p^*
10^3
10^1
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-9}
f($x^{(k)}$) - p^*
0
250
500
750
1000
1250
1500
1750
Iteration

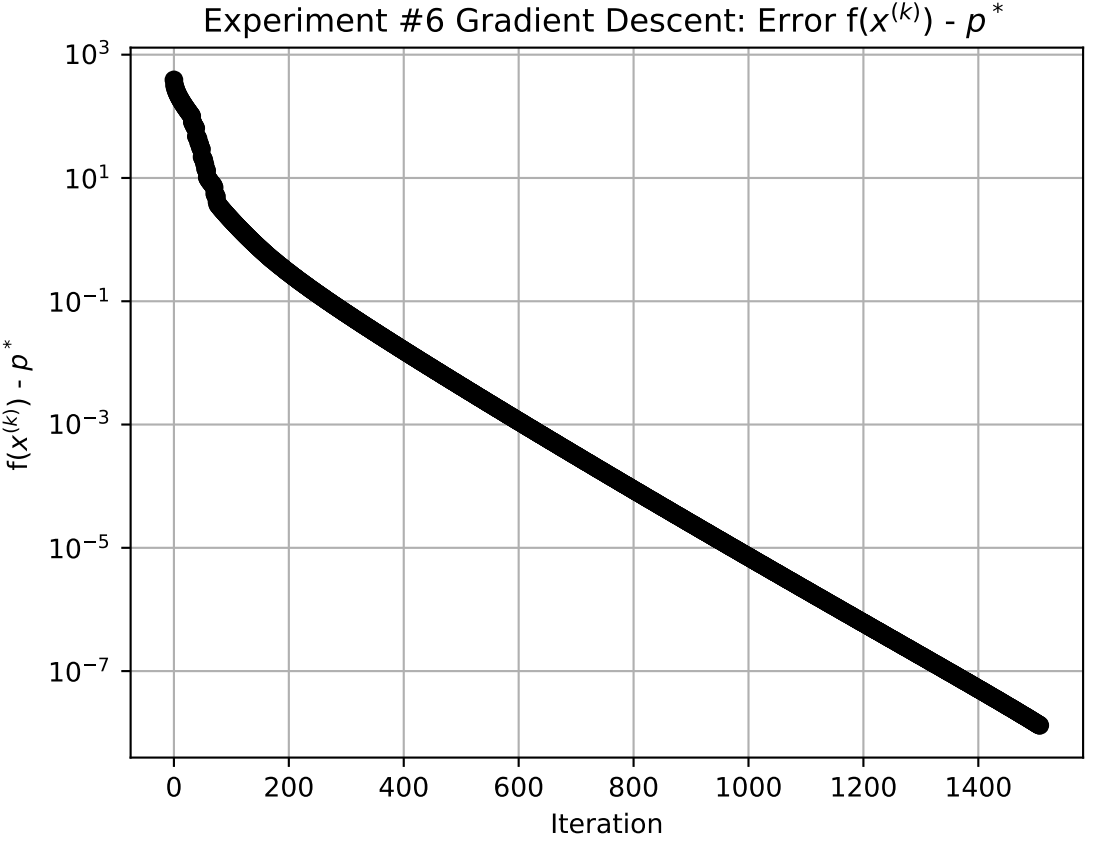

Experiment #4 Gradient Descent: Step Size
Step Size
10^{-2}
10^{-3}
10^{-4}
0
250
500
750
1000
1250
1500
1750
Iteration

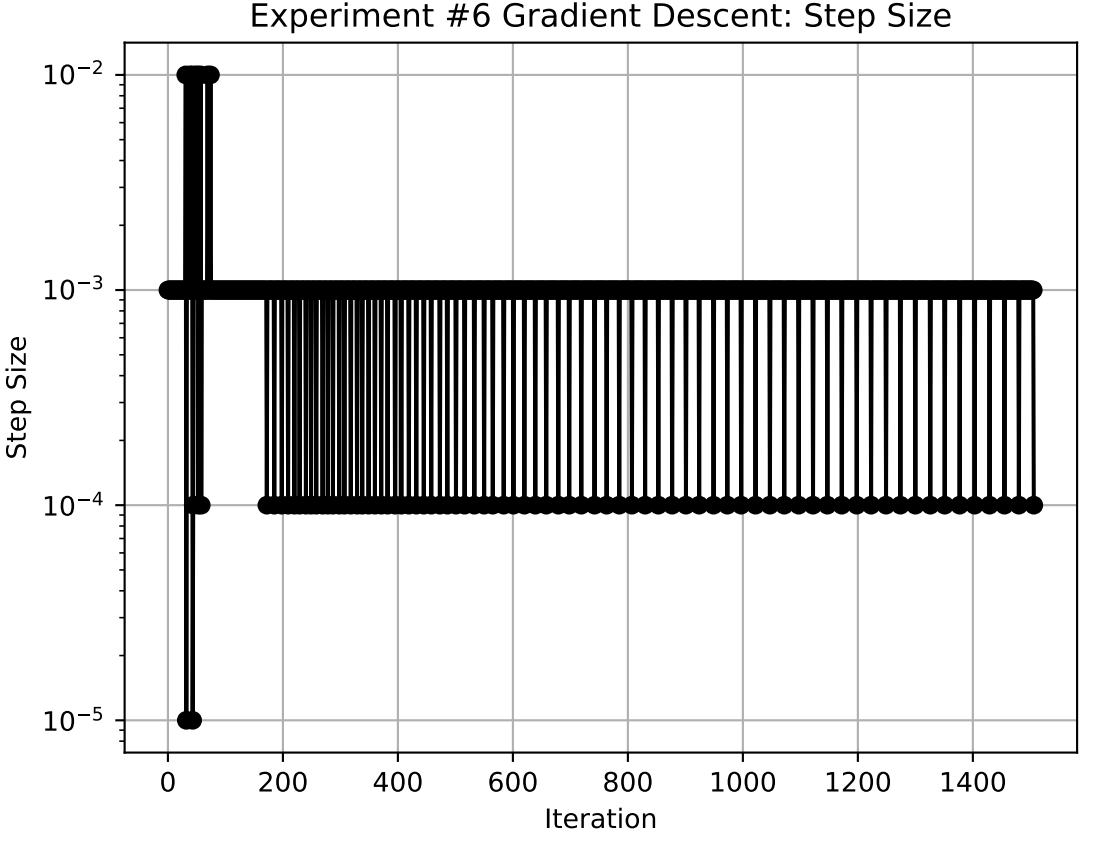

Experiment #5 Gradient Descent: f($x^{(k)}$)
0
−50
−100
−150
−200
−250
−300
−350
−400
f($x^{(k)}$)
0
200
400
600
800
1000
1200
1400
Iteration

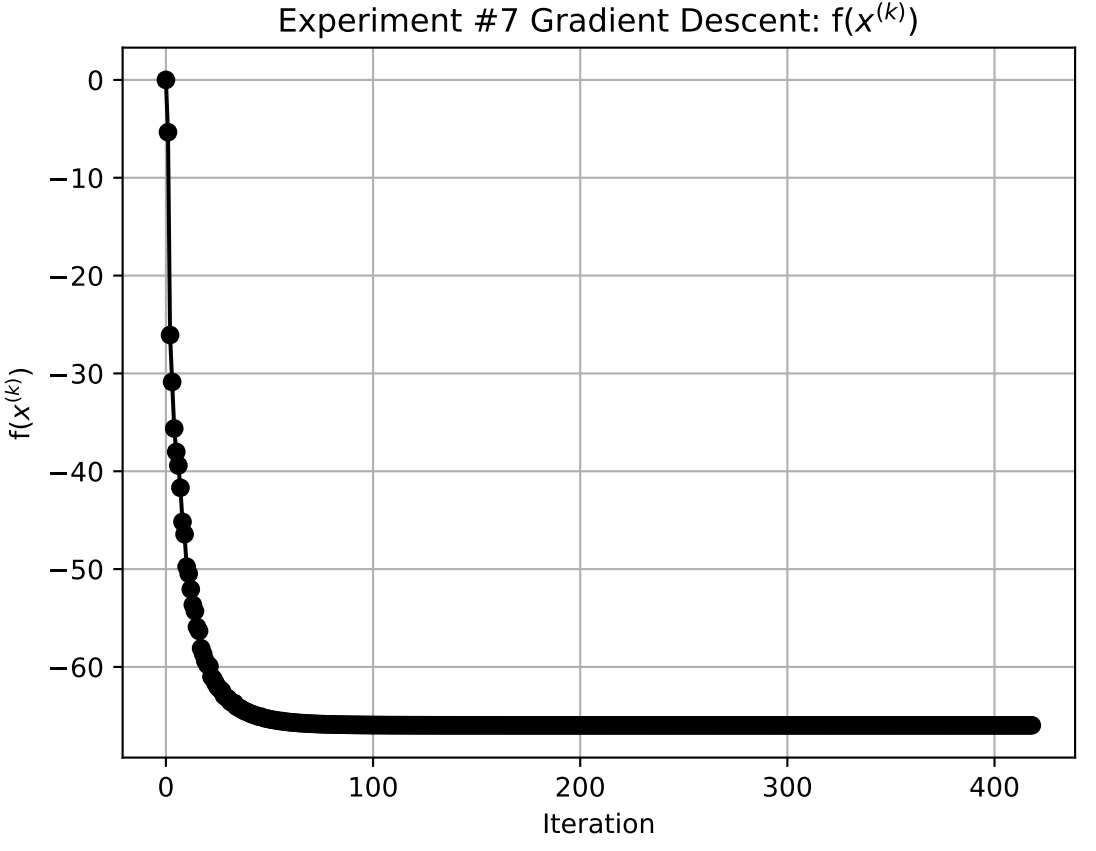

Experiment #5 Gradient Descent: Error f($x^{(k)}$) - p^*
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
f($x^{(k)}$) - p^*
0
200
400
600
800
1000
1200
1400
Iteration

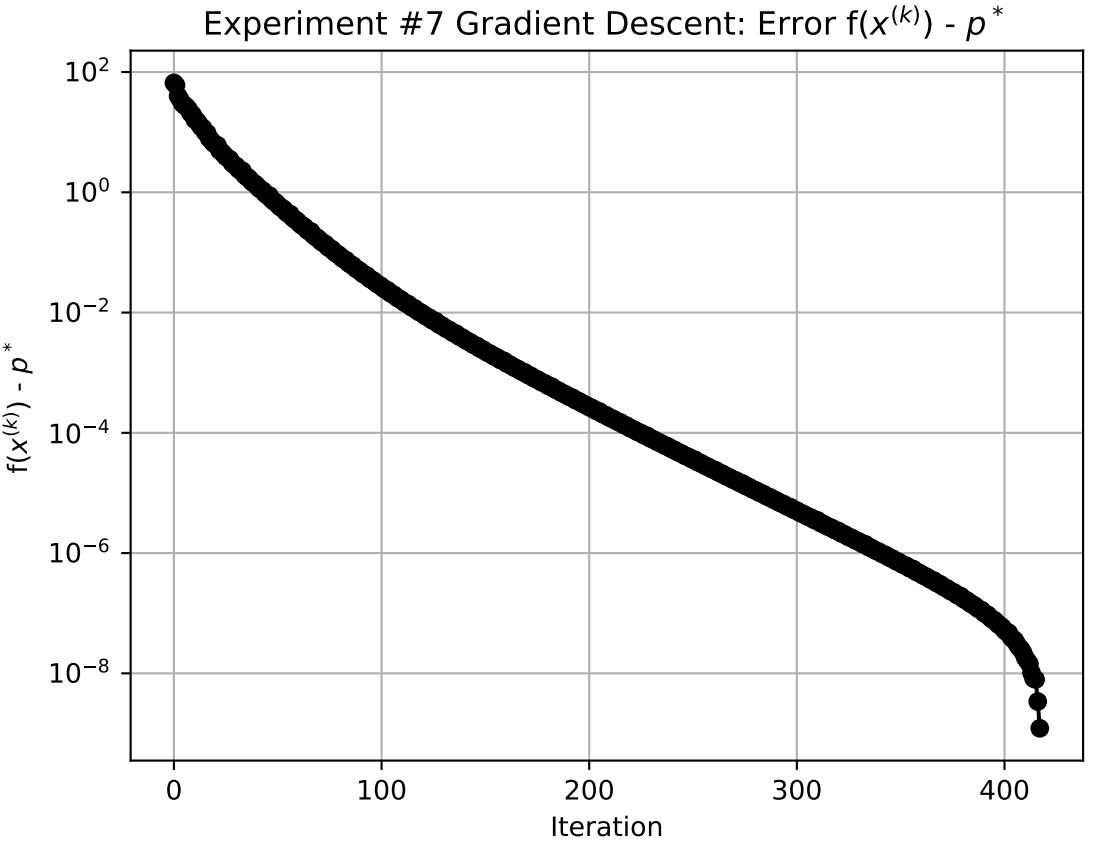

Experiment #5 Gradient Descent: Step Size
Step Size
10^{-3}
0
200
400
600
800
1000
1200
1400
Iteration

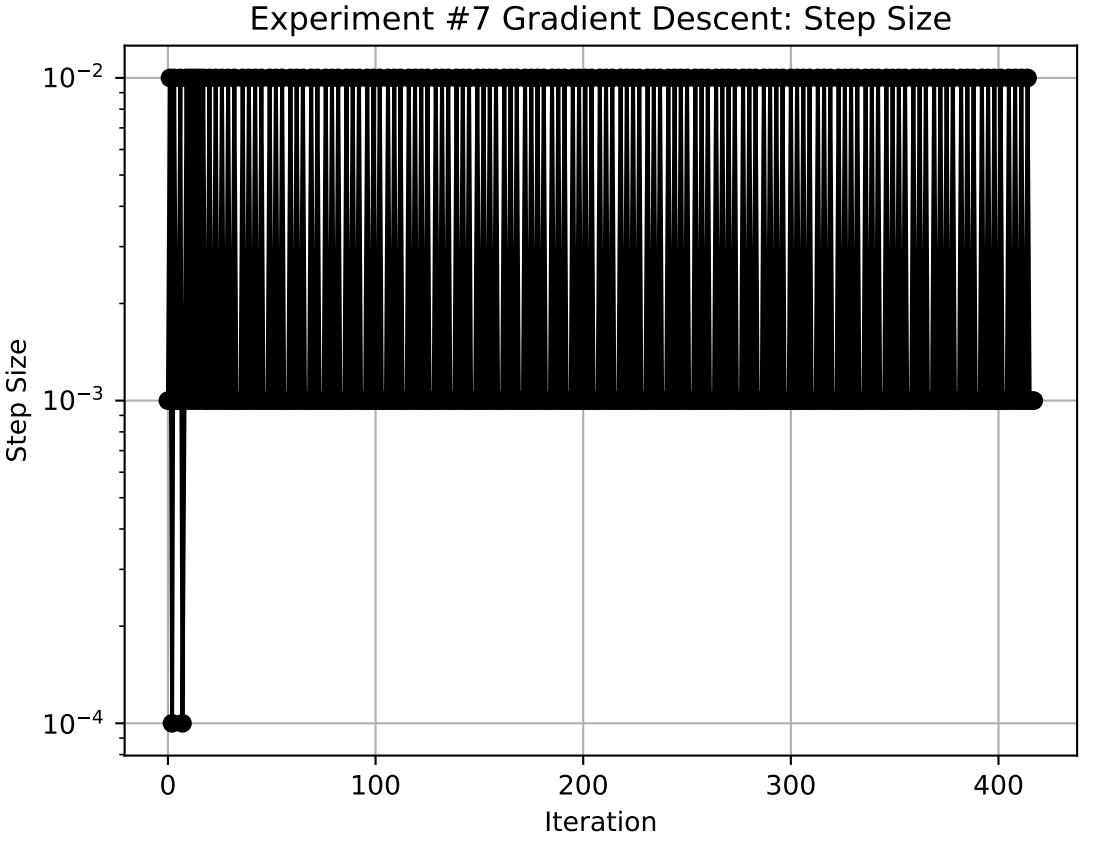

Experiment #6 Gradient Descent: f($x^{(k)}$)
0
−50
−100
−150
−200
−250
−300
−350
−400
f($x^{(k)}$)
0
200
400
600
800
1000
1200
1400
Iteration

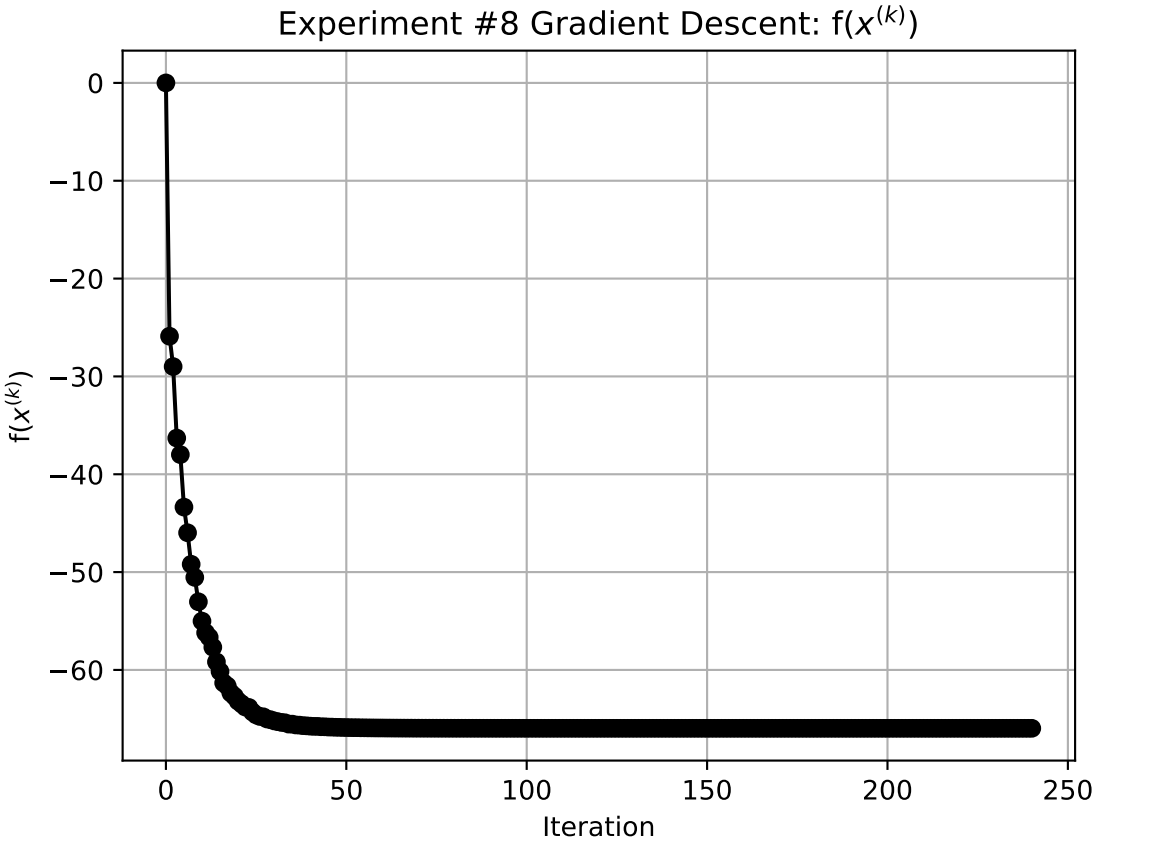

Experiment #6 Gradient Descent: Error f($x^{(k)}$) - p^*
10^3
10^1
10^{-1}
10^{-3}
10^{-5}
10^{-7}
f($x^{(k)}$) - p^*
0
200
400
600
800
1000
1200
1400
Iteration

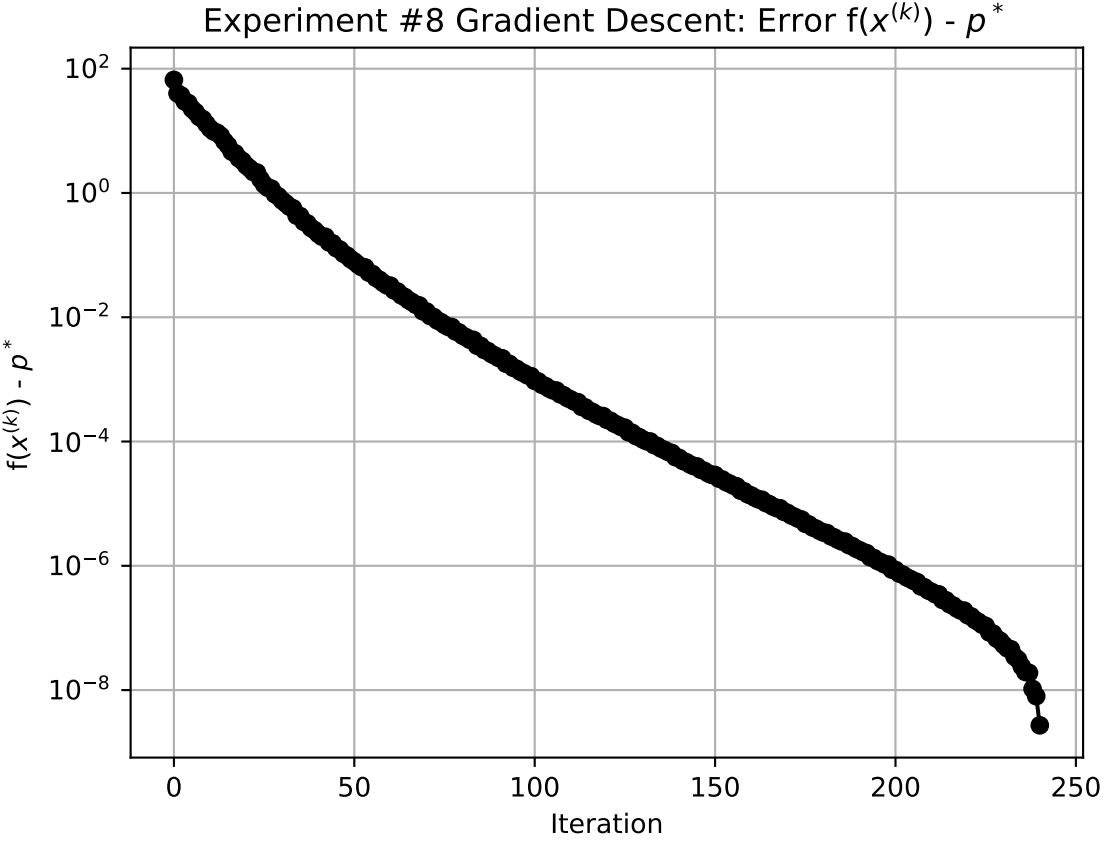

Experiment #6 Gradient Descent: Step Size
Step Size
10^{-2}
10^{-3}
10^{-4}
10^{-5}
0
200
400
600
800
1000
1200
1400
Iteration

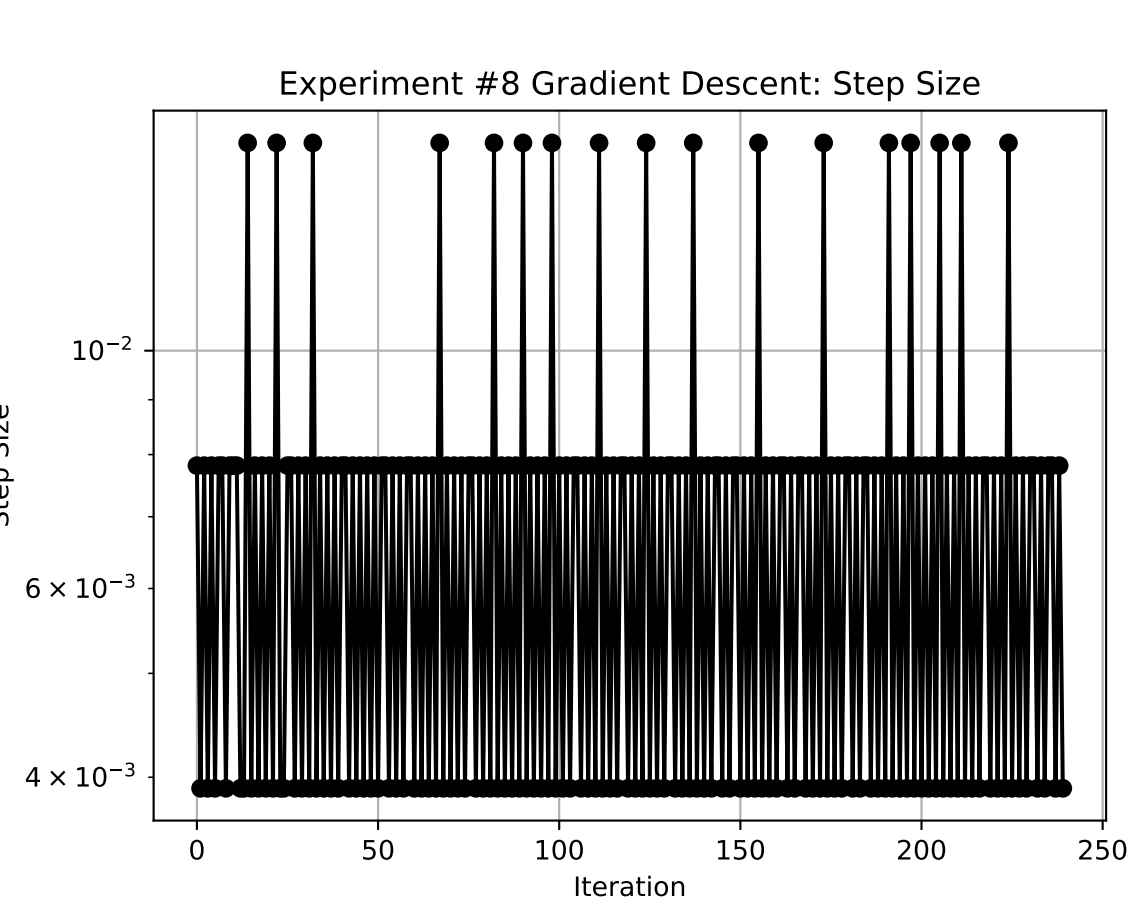

Experiment #7 Gradient Descent: f($x^{(k)}$)
0
−10
−20
−30
−40
−50
−60
f($x^{(k)}$)
0
100
200
300
400
Iteration

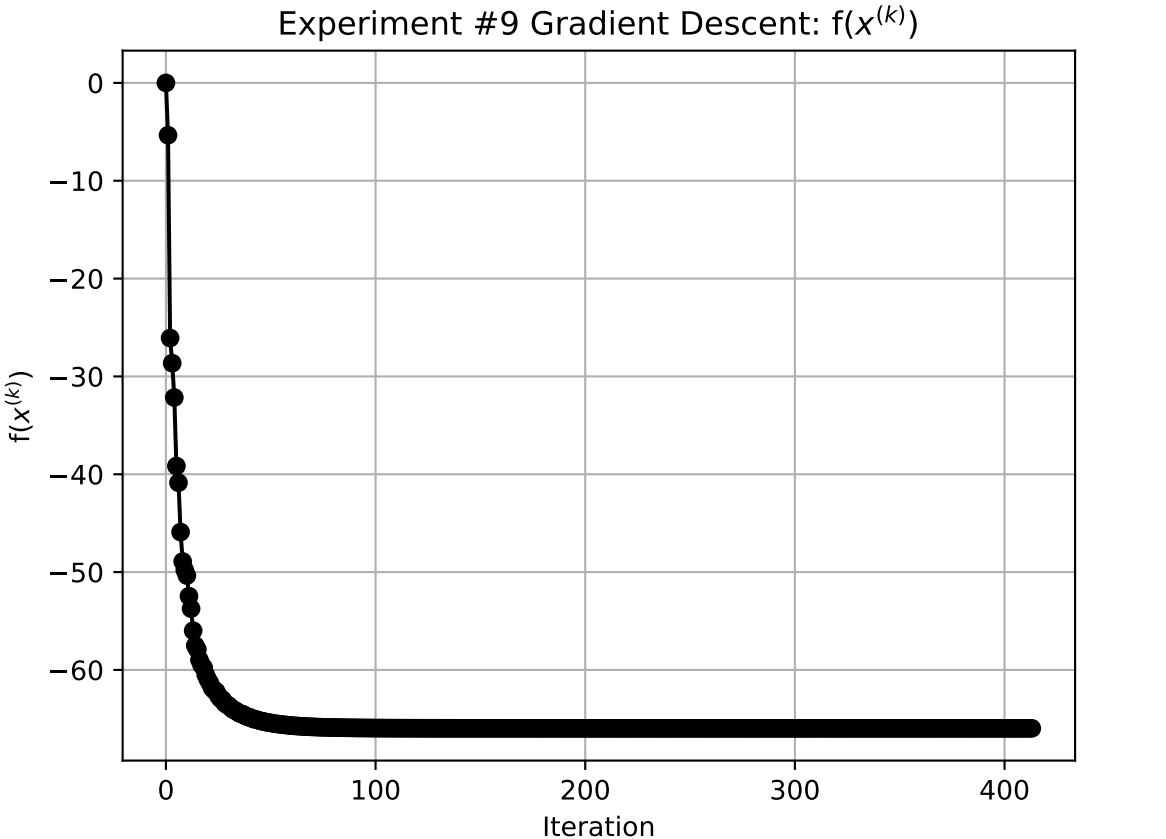

Experiment #7 Gradient Descent: Error f($x^{(k)}$) - p^*
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
f($x^{(k)}$) - p^*
0
100
200
300
400
Iteration


Experiment #7 Gradient Descent: Step Size
Step Size
10^{-2}
10^{-3}
10^{-4}
0
100
200
300
400
Iteration

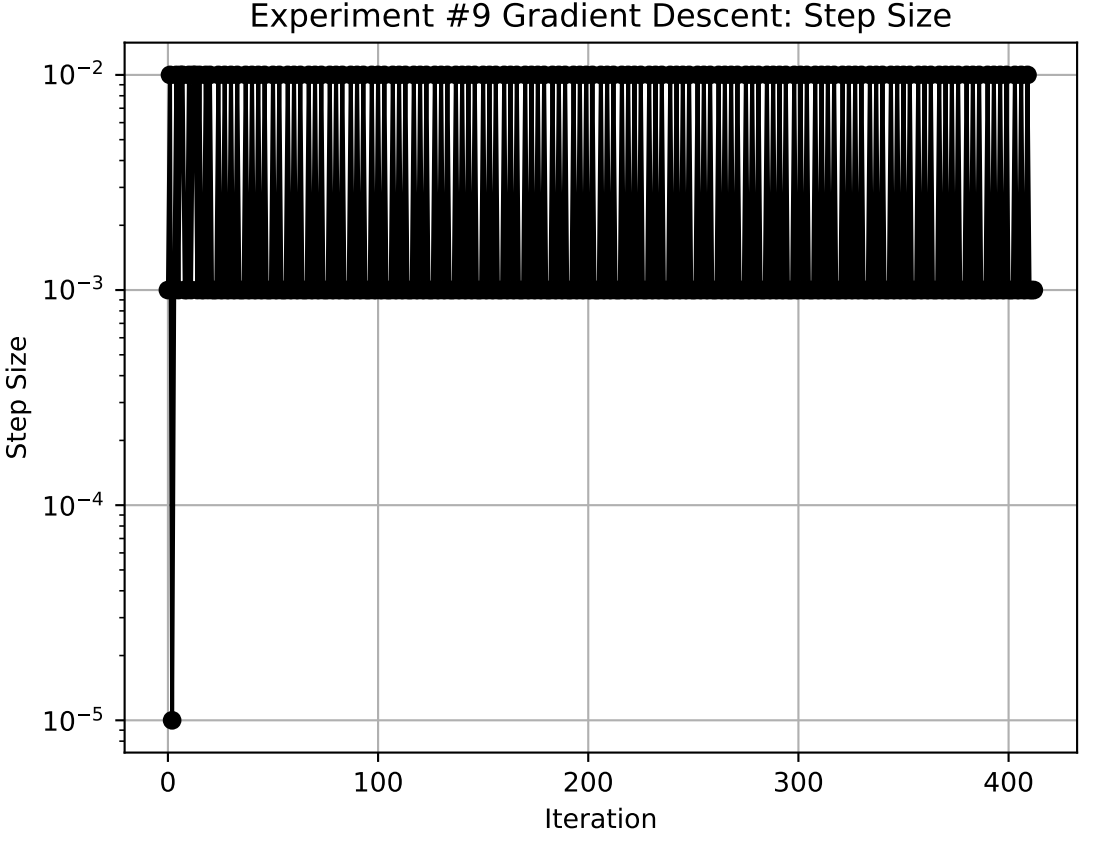

Experiment #8 Gradient Descent: f($x^{(k)}$)
0
−10
−20
−30
−40
−50
−60
f($x^{(k)}$)
0
50
100
150
200
250
Iteration

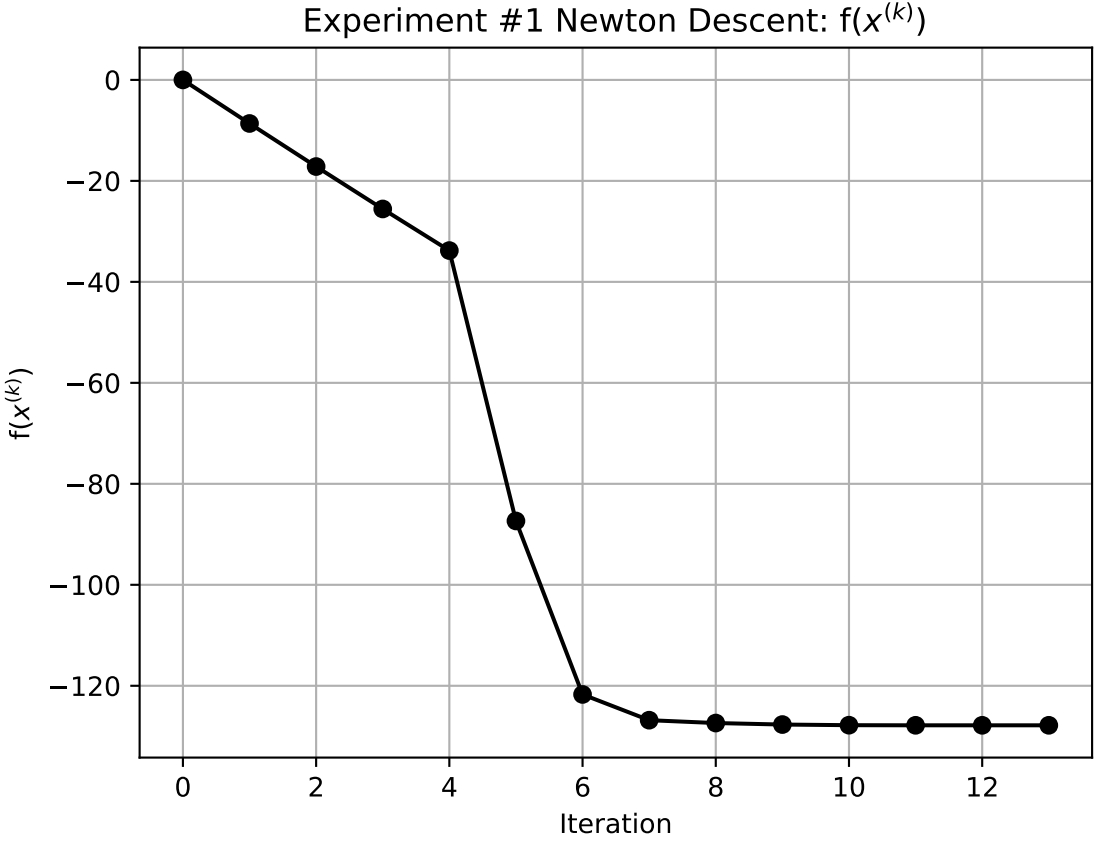

Experiment #8 Gradient Descent: Error f($x^{(k)}$) - p^*
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
f($x^{(k)}$) - p^*
0
50
100
150
200
250
Iteration

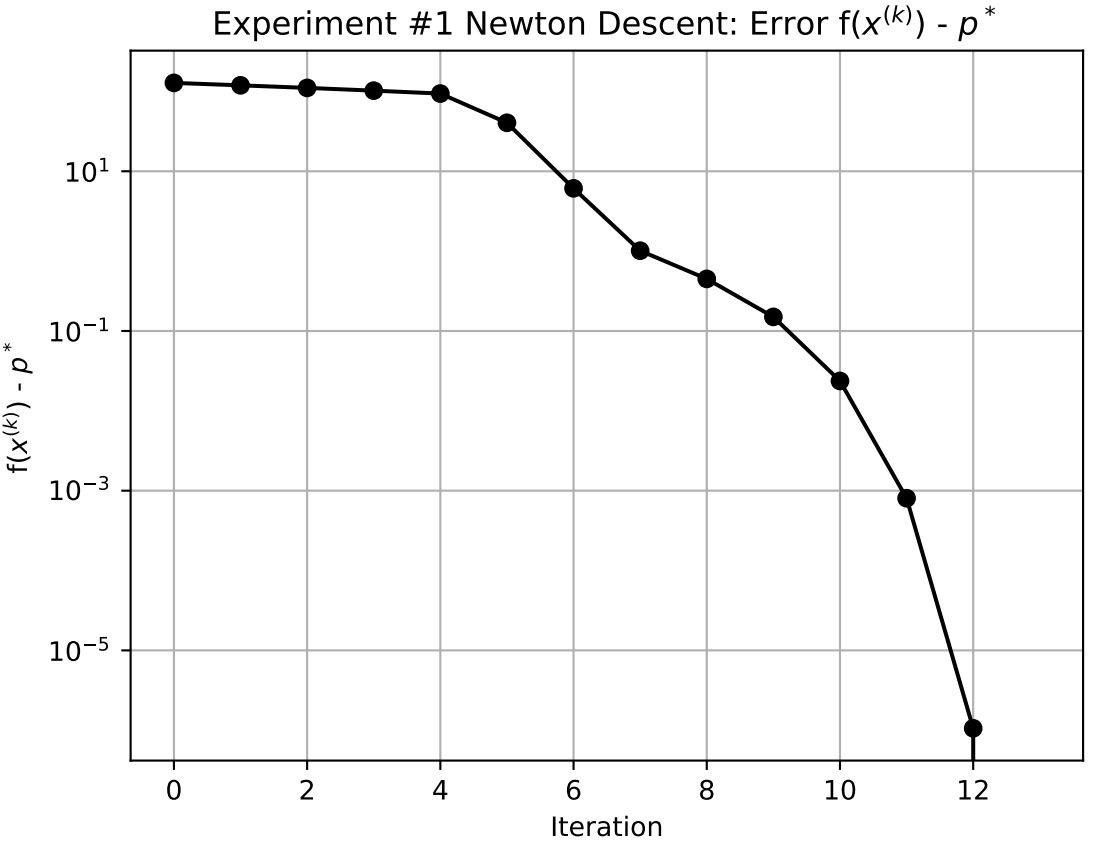

Experiment #8 Gradient Descent: Step Size
Step Size
10^{-2}
6×10^{-3}
4×10^{-3}
0
50
100
150
200
250
Iteration

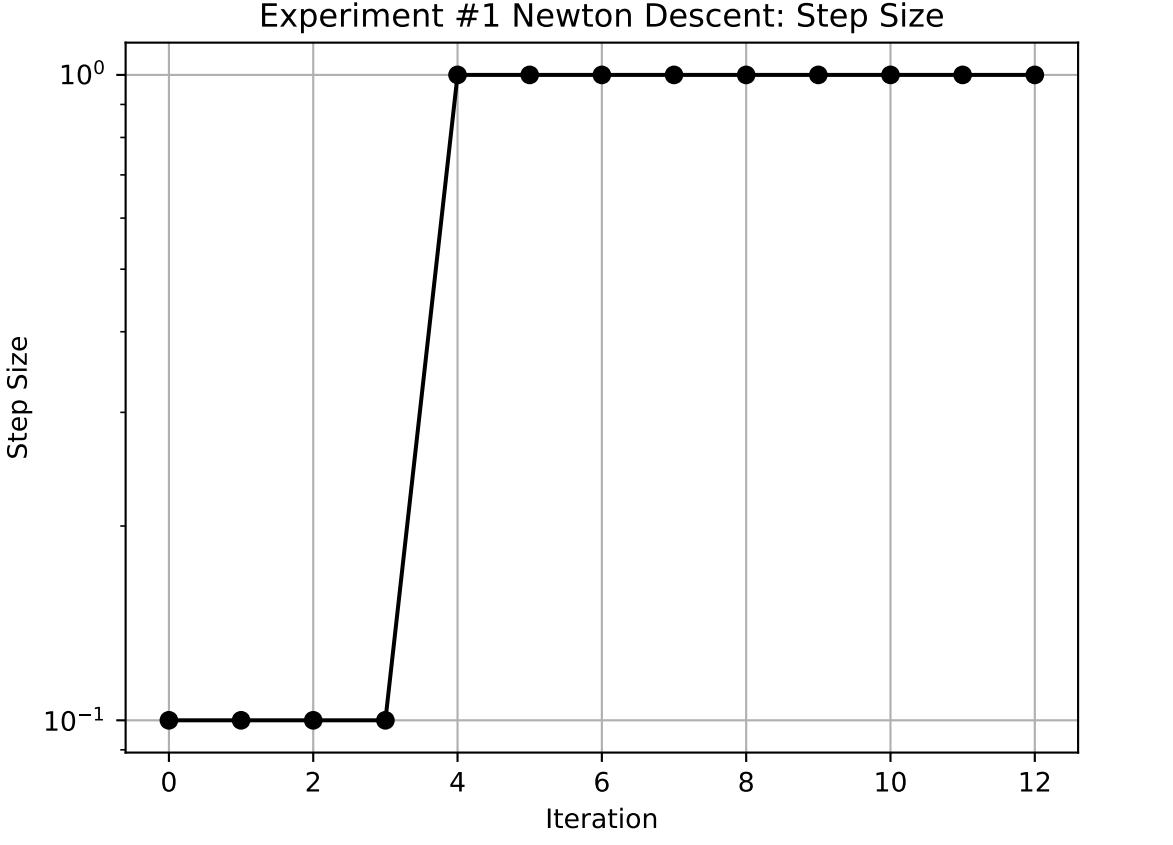

Experiment #9 Gradient Descent: f($x^{(k)}$)
0
−10
−20
−30
−40
−50
−60
f($x^{(k)}$)
0
100
200
300
400
Iteration

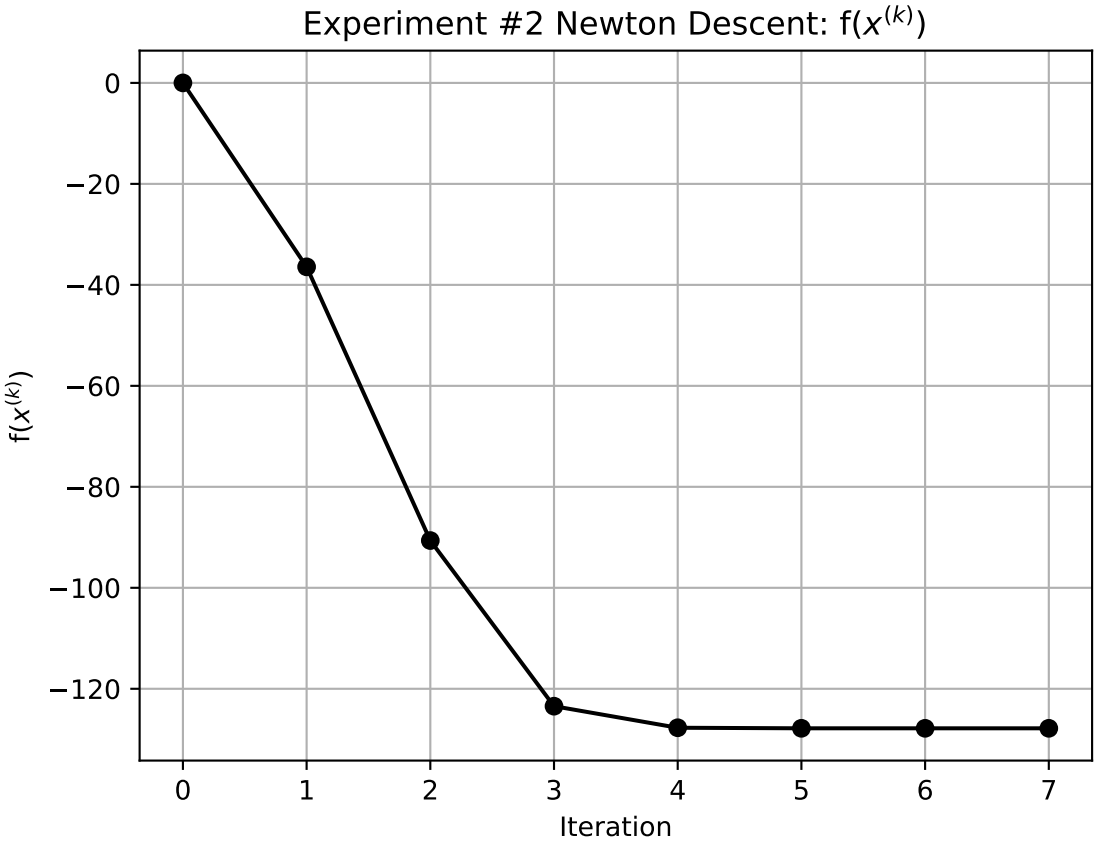

Experiment #9 Gradient Descent: Error f($x^{(k)}$) - p^*
10^1
10^{-1}
10^{-3}
10^{-5}
10^{-7}
f($x^{(k)}$) - p^*
0
100
200
300
400
Iteration

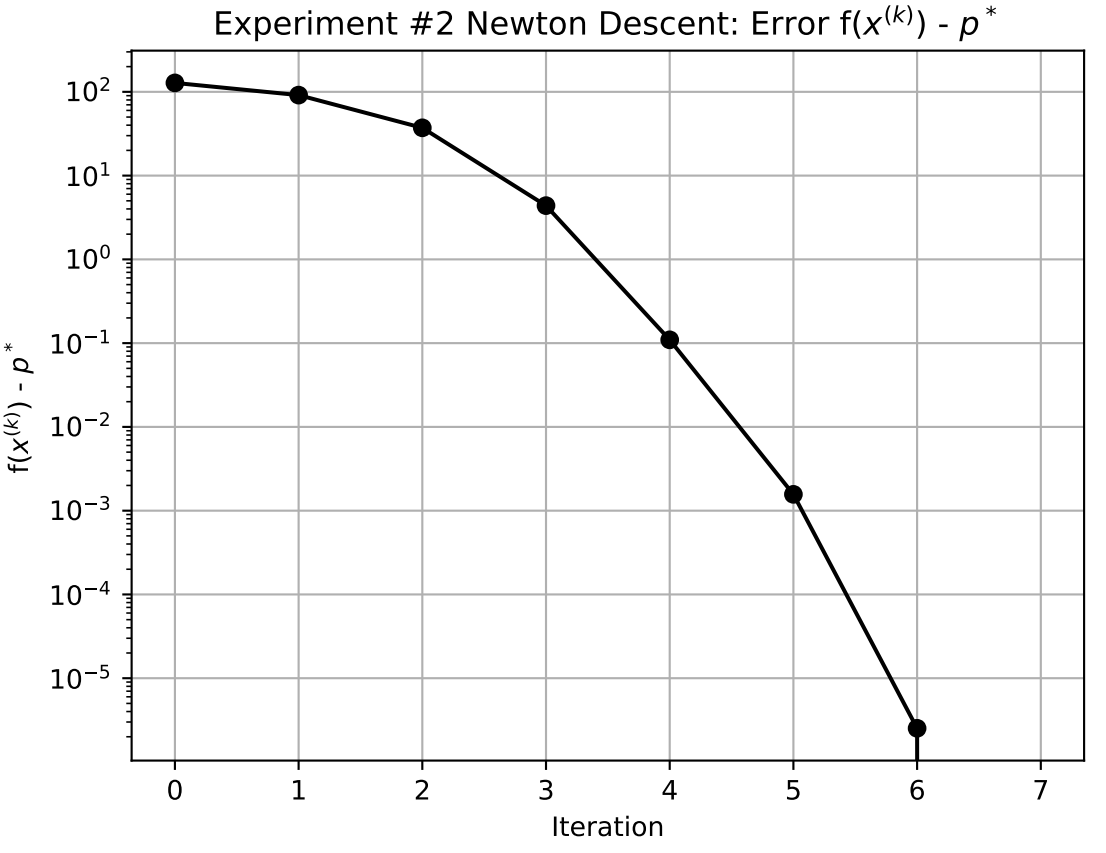

Experiment #9 Gradient Descent: Step Size
Step Size
10^{-2}
10^{-3}
10^{-4}
10^{-5}
0
100
200
300
400
Iteration

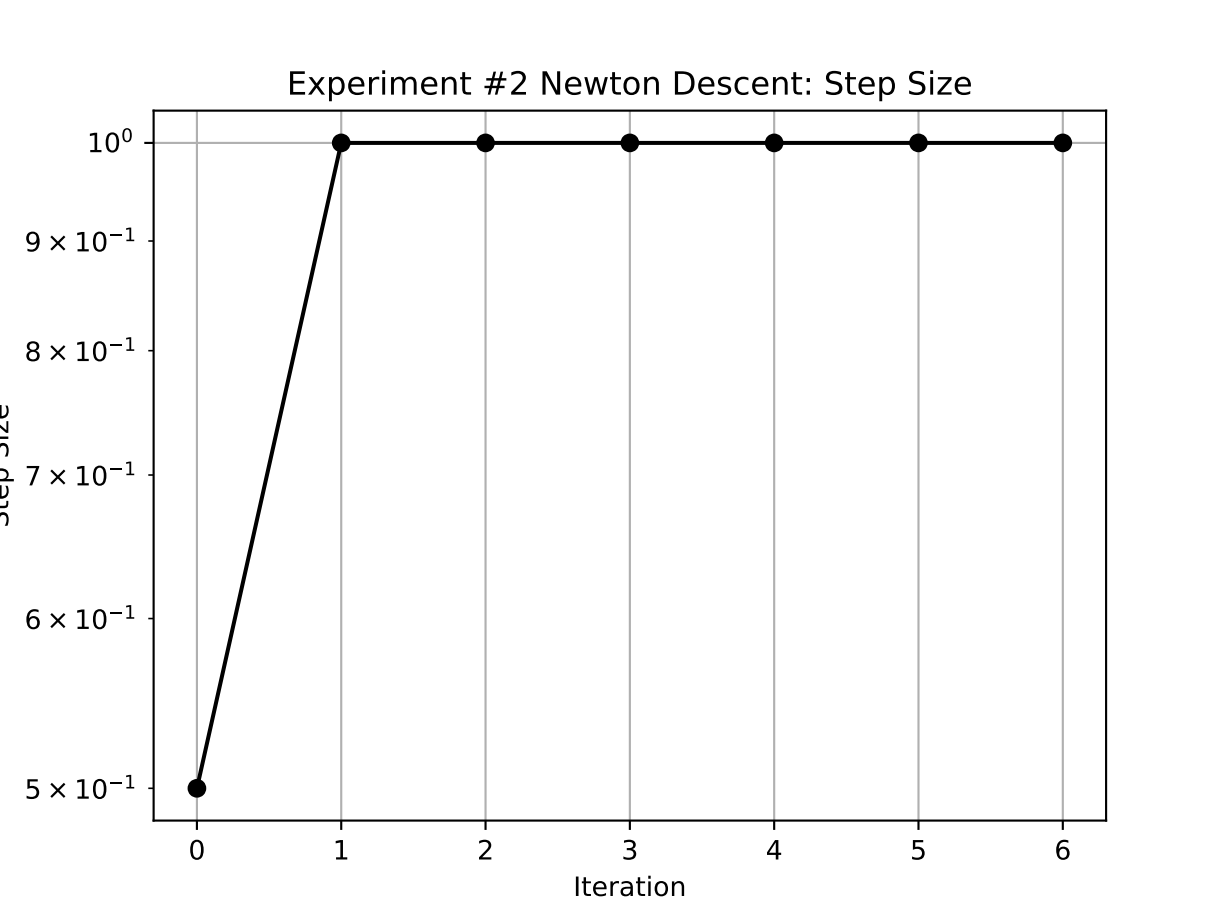

Experiment #1 Newton Descent: f($x^{(k)}$)
0
−20
−40
−60
−80
−100
−120
f($x^{(k)}$)
0
2
4
6
8
10
12
Iteration

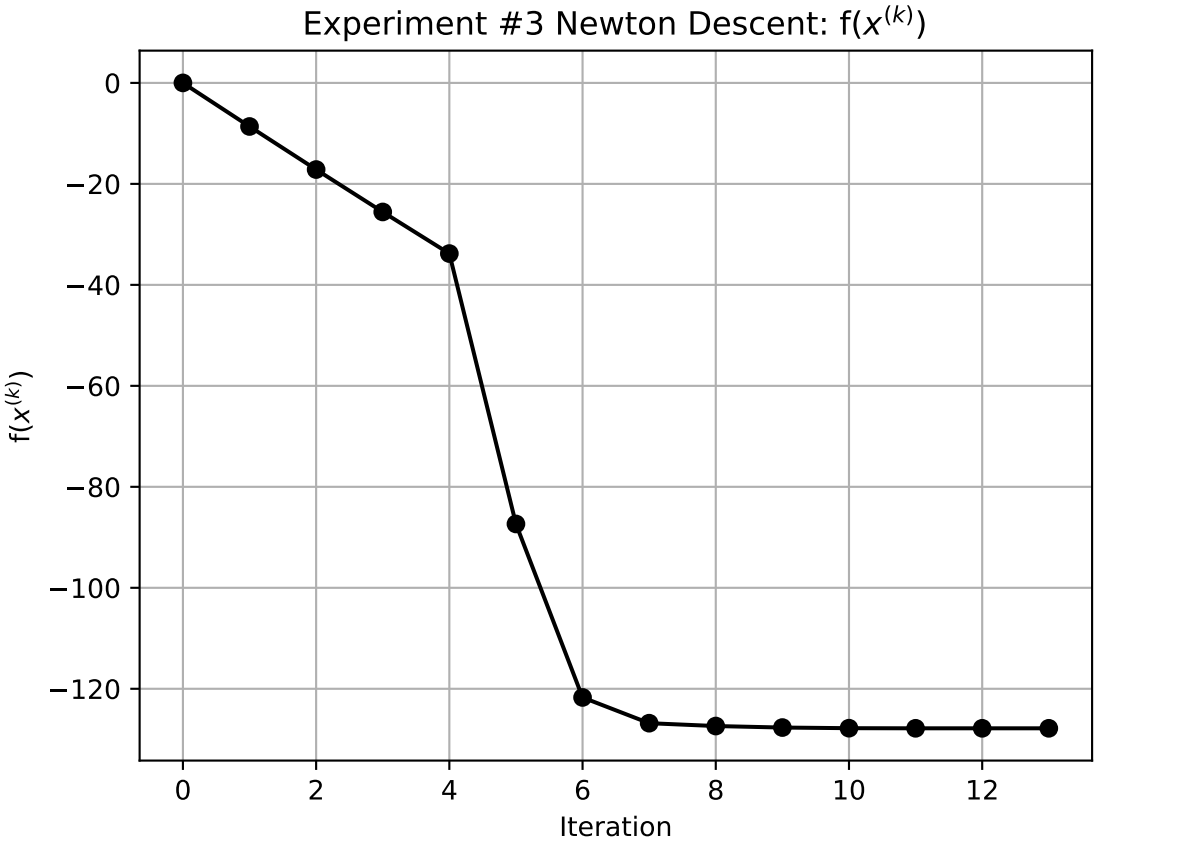

Experiment #1 Newton Descent: Error f($x^{(k)}$) - p^*
10^1
10^{-1}
10^{-3}
10^{-5}
f($x^{(k)}$) - p^*
0
2
4
6
8
10
12
Iteration

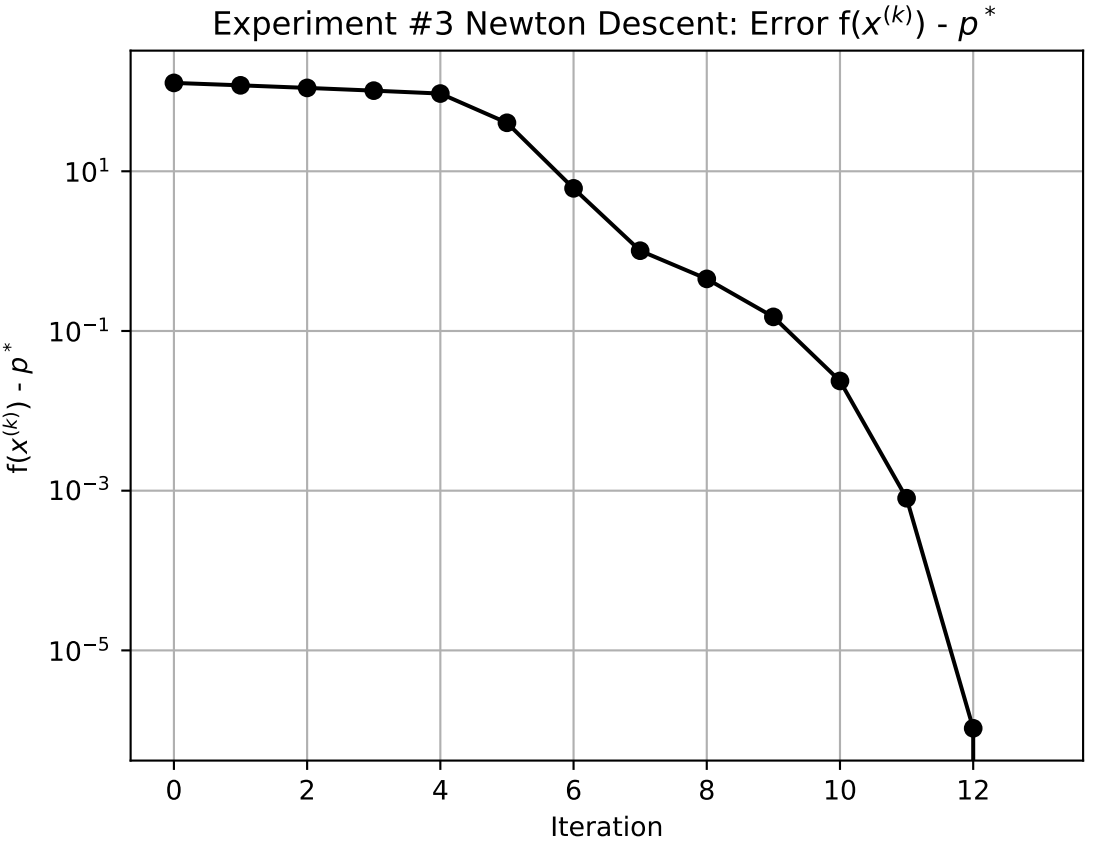

Experiment #1 Newton Descent: Step Size
Step Size
10^0
10^{-1}
0
2
4
6
8
10
12
Iteration

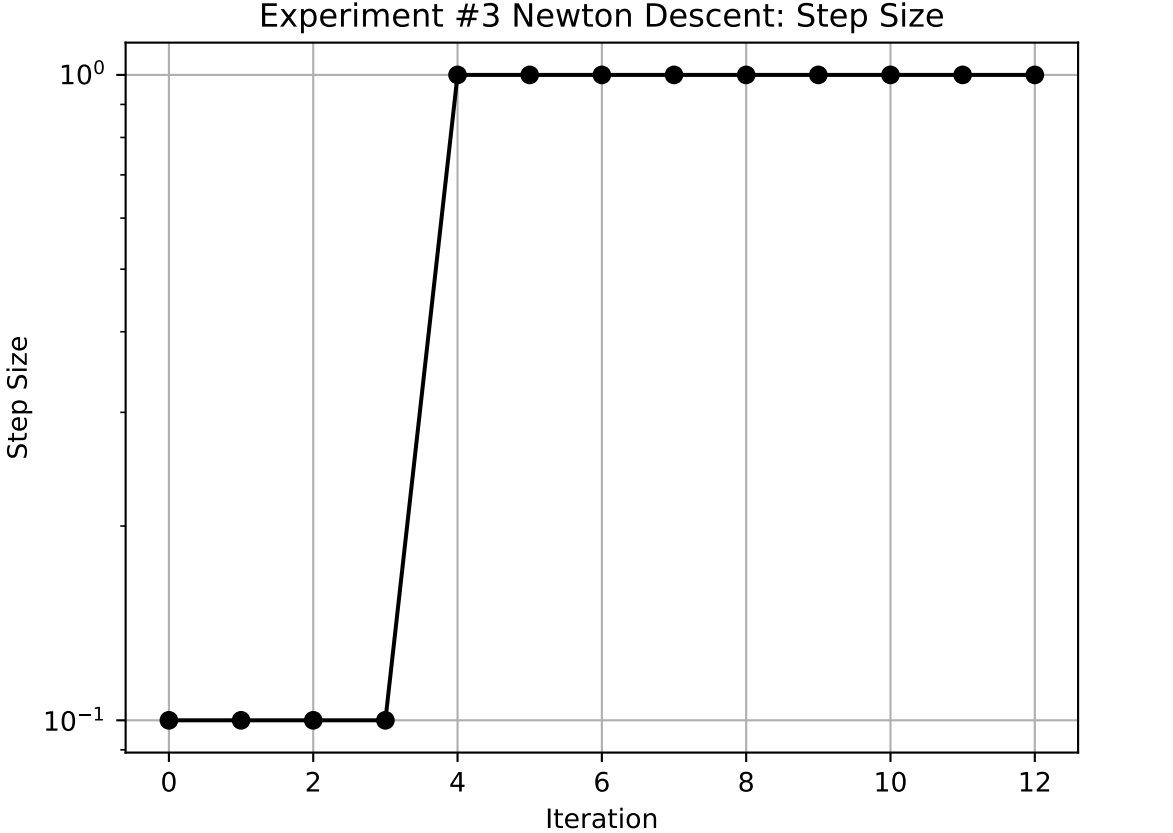

Experiment #2 Newton Descent: f($x^{(k)}$)
0
−20
−40
−60
−80
−100
−120
f($x^{(k)}$)
0
1
2
3
4
5
6
7
Iteration

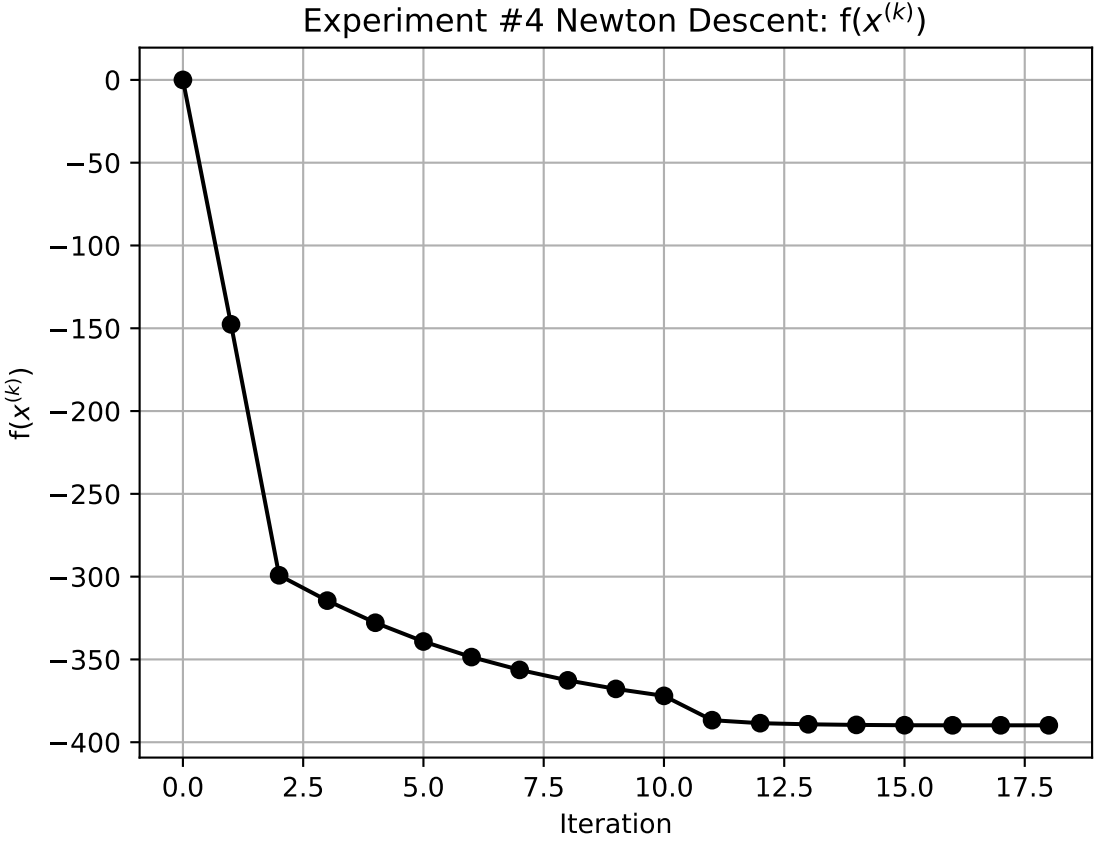

Experiment #2 Newton Descent: Error f($x^{(k)}$) - p^*
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
f($x^{(k)}$) - p^*
0
1
2
3
4
5
6
7
Iteration

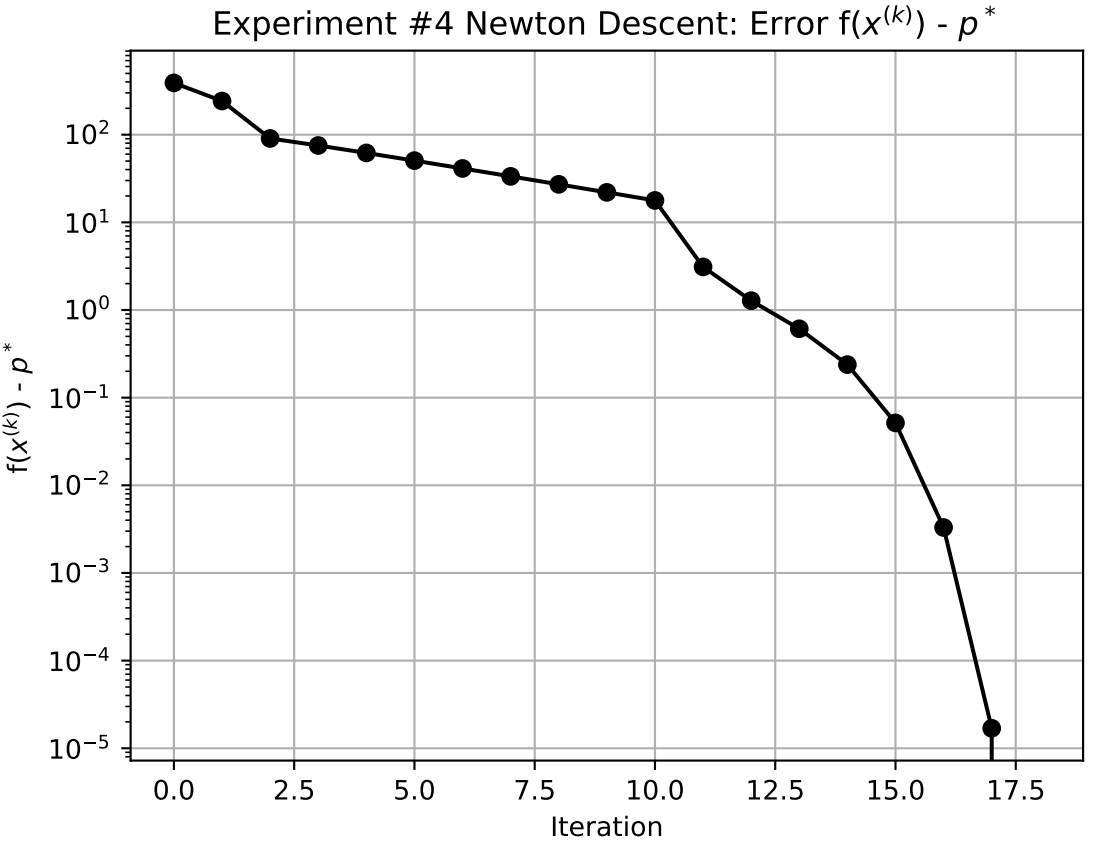

Experiment #2 Newton Descent: Step Size
Step Size
Iteration
10^0
9×10^{-1}
8×10^{-1}
7×10^{-1}
6×10^{-1}
5×10^{-1}
0
1
2
3
4
5
6

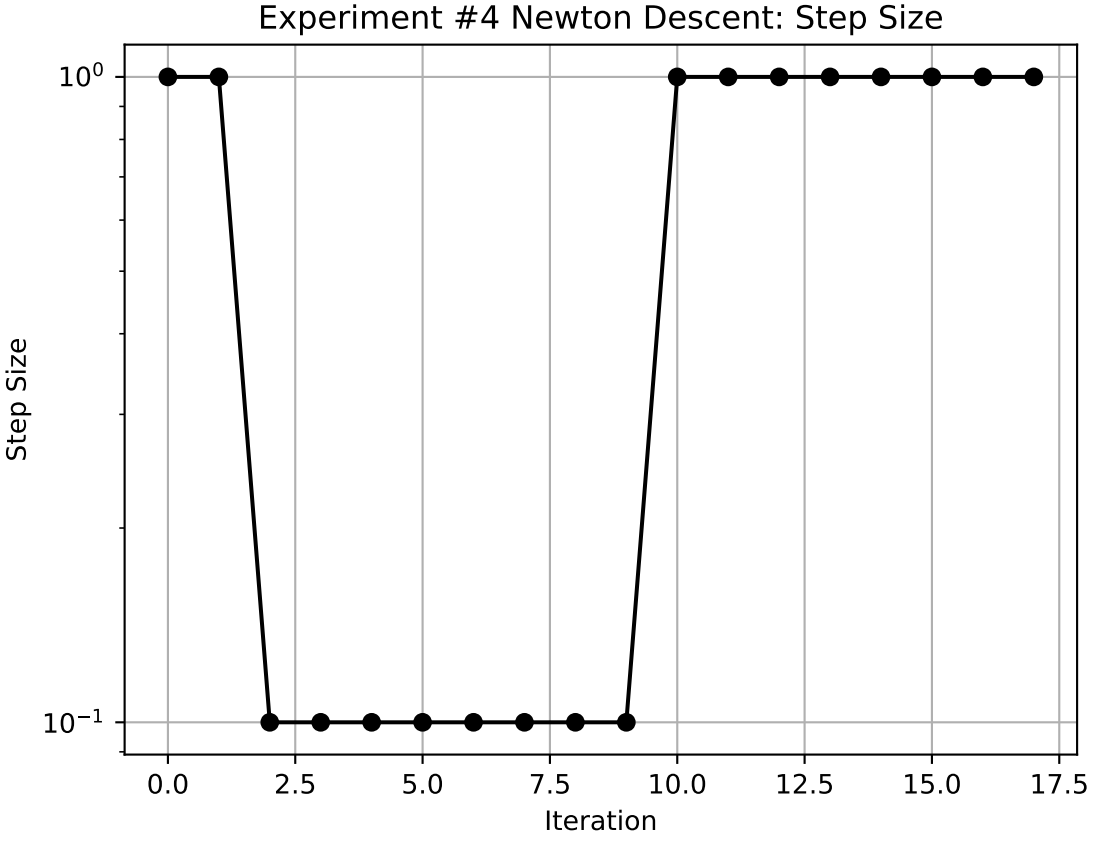

Experiment #3 Newton Descent: f($x^{(k)}$)
0
−20
−40
−60
−80
−100
−120
f($x^{(k)}$)
0
2
4
6
8
10
12
Iteration

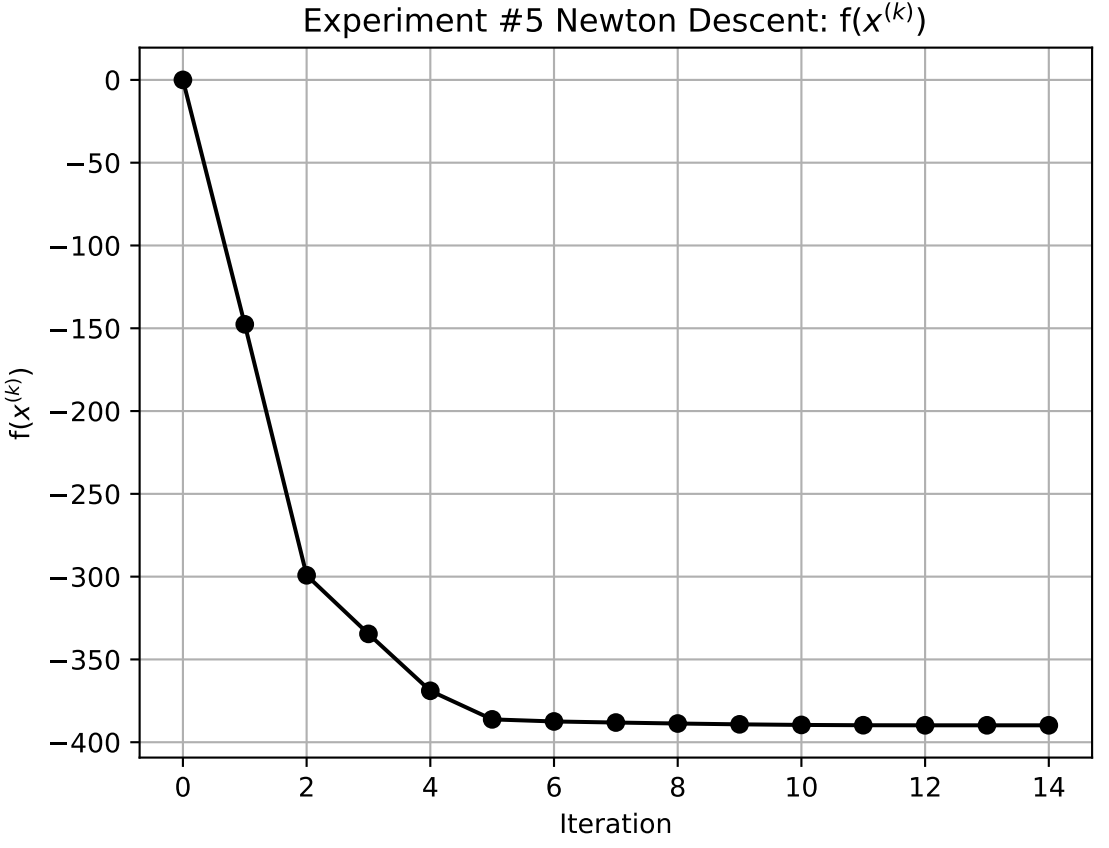

Experiment #3 Newton Descent: Error f($x^{(k)}$) - p^*
10^1
10^{-1}
10^{-3}
10^{-5}
f($x^{(k)}$) - p^*
0
2
4
6
8
10
12
Iteration

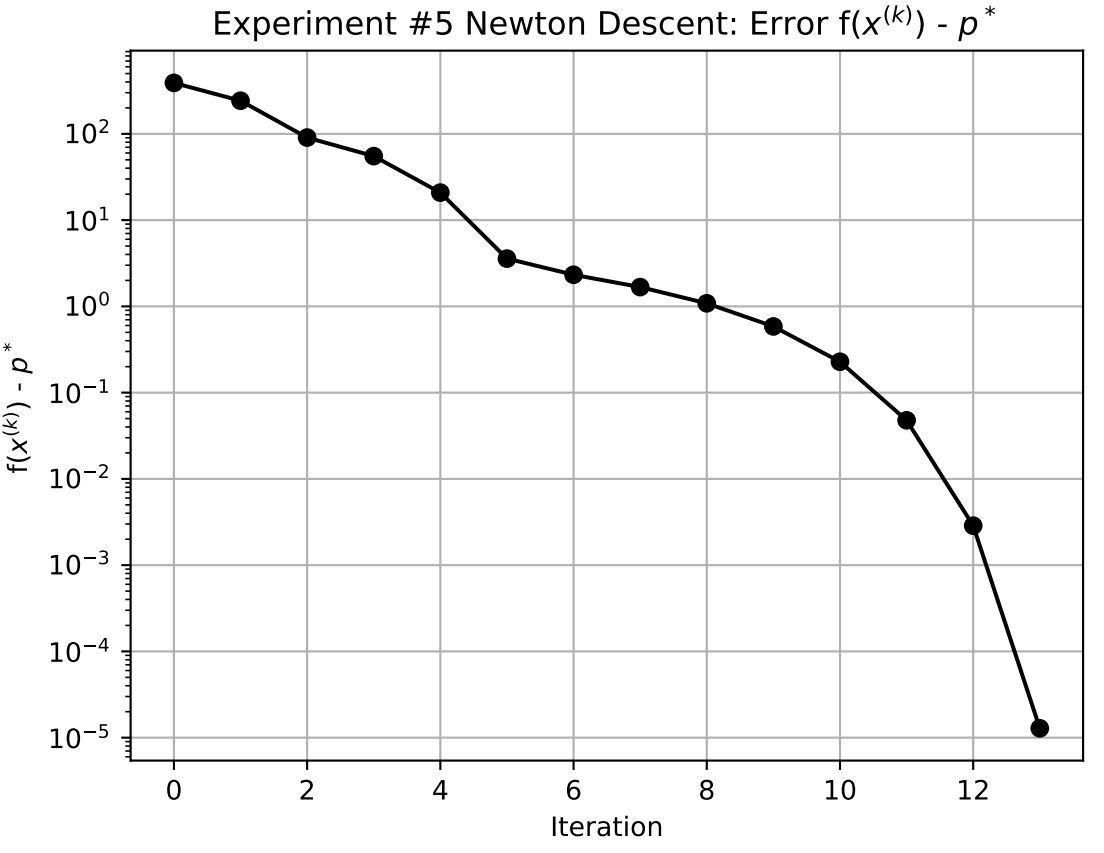

Experiment #3 Newton Descent: Step Size
10^0
10^{-1}
Step Size
0
2
4
6
8
10
12
Iteration

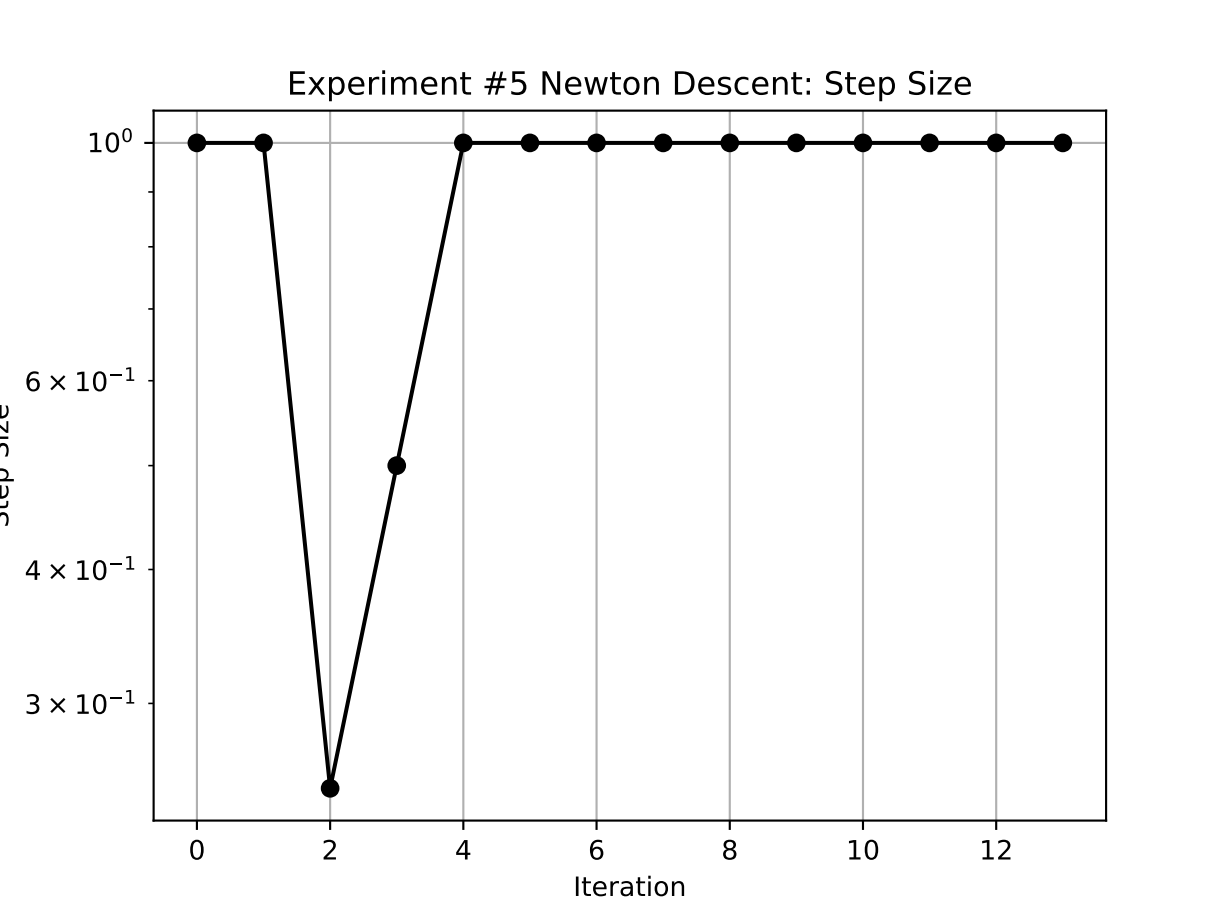

Experiment #4 Newton Descent: f($x^{(k)}$)
0
−50
−100
−150
−200
−250
−300
−350
−400
f($x^{(k)}$)
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Iteration

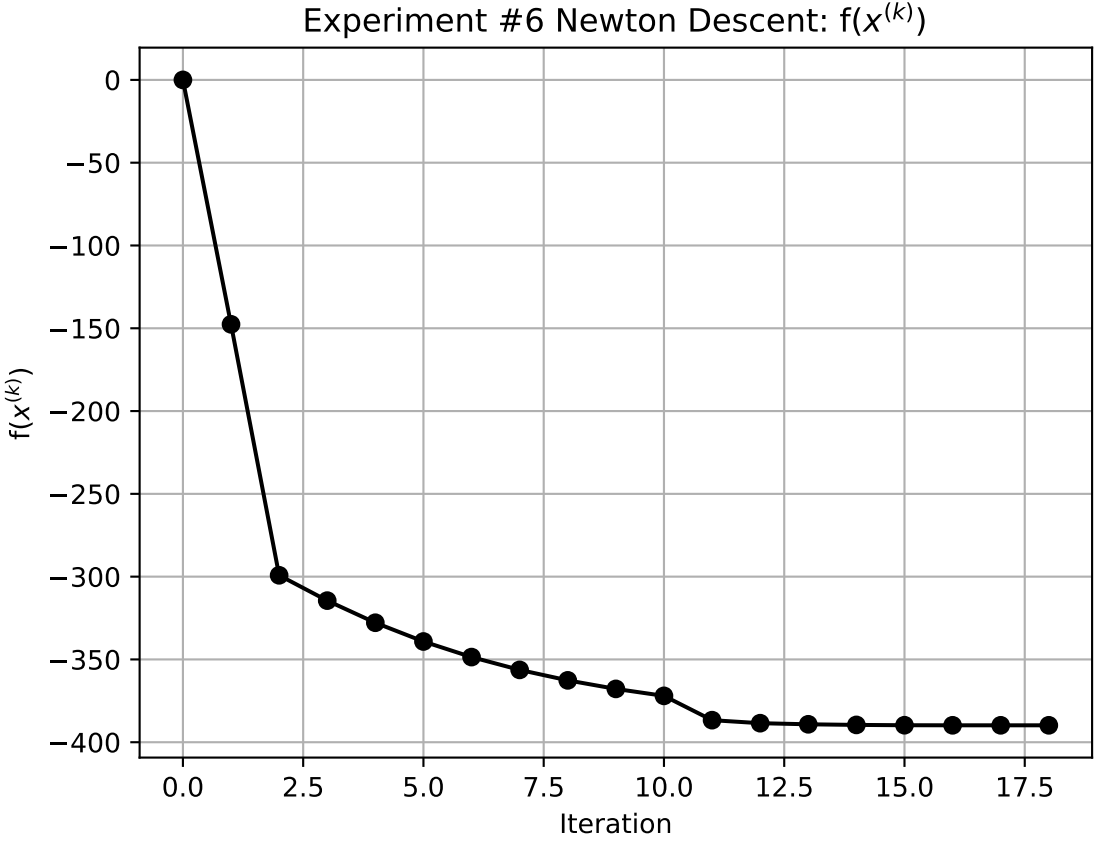

Experiment #4 Newton Descent: Error f($x^{(k)}$) - p^*
f($x^{(k)}$) - p^*
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Iteration

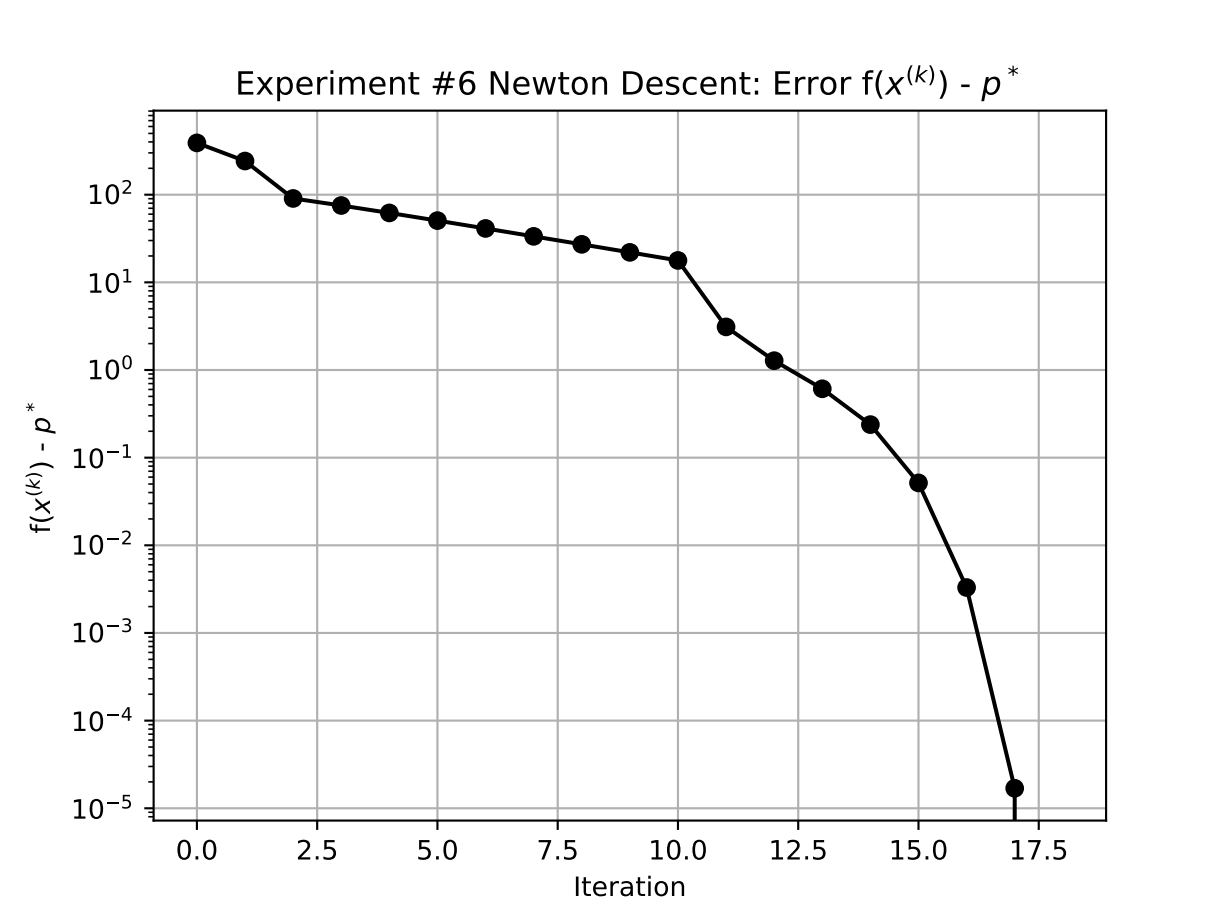

Experiment #4 Newton Descent: Step Size
10^0
10^{-1}
Step Size
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Iteration

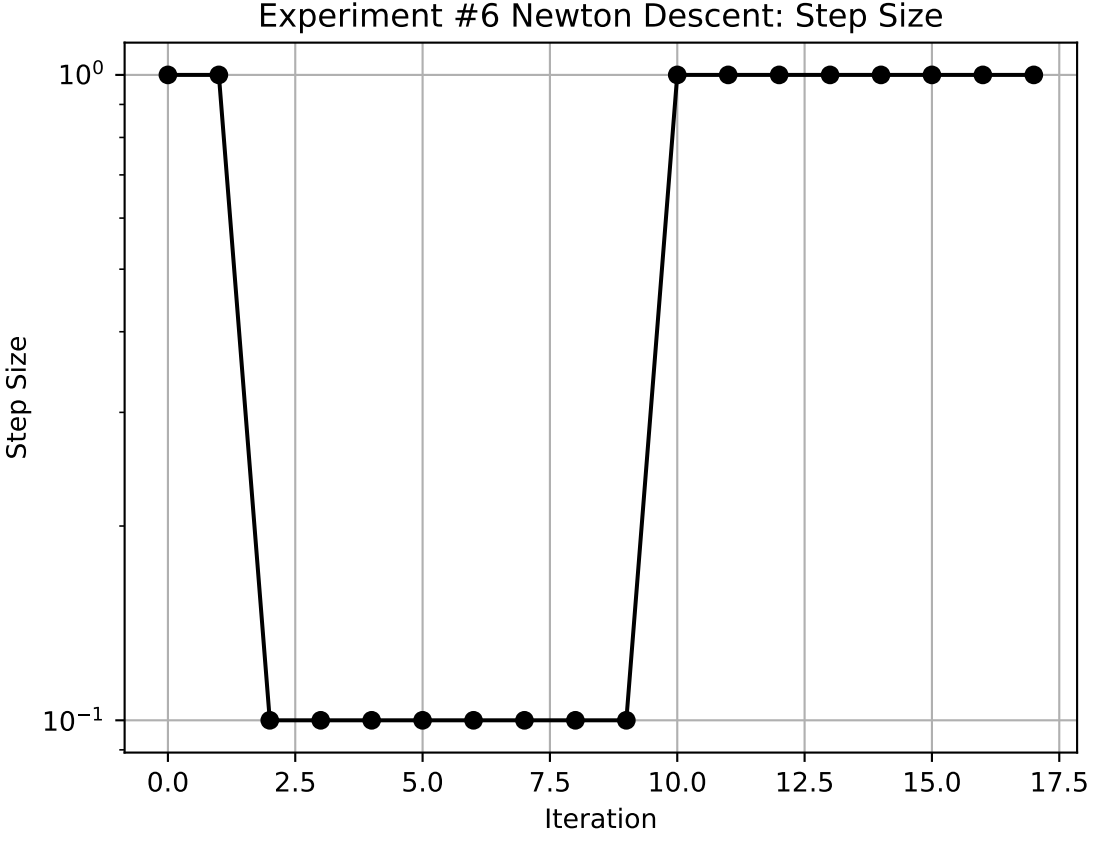

Experiment #5 Newton Descent: f($x^{(k)}$)
0
−50
−100
−150
−200
−250
−300
−350
−400
f($x^{(k)}$)
0
2
4
6
8
10
12
14
Iteration

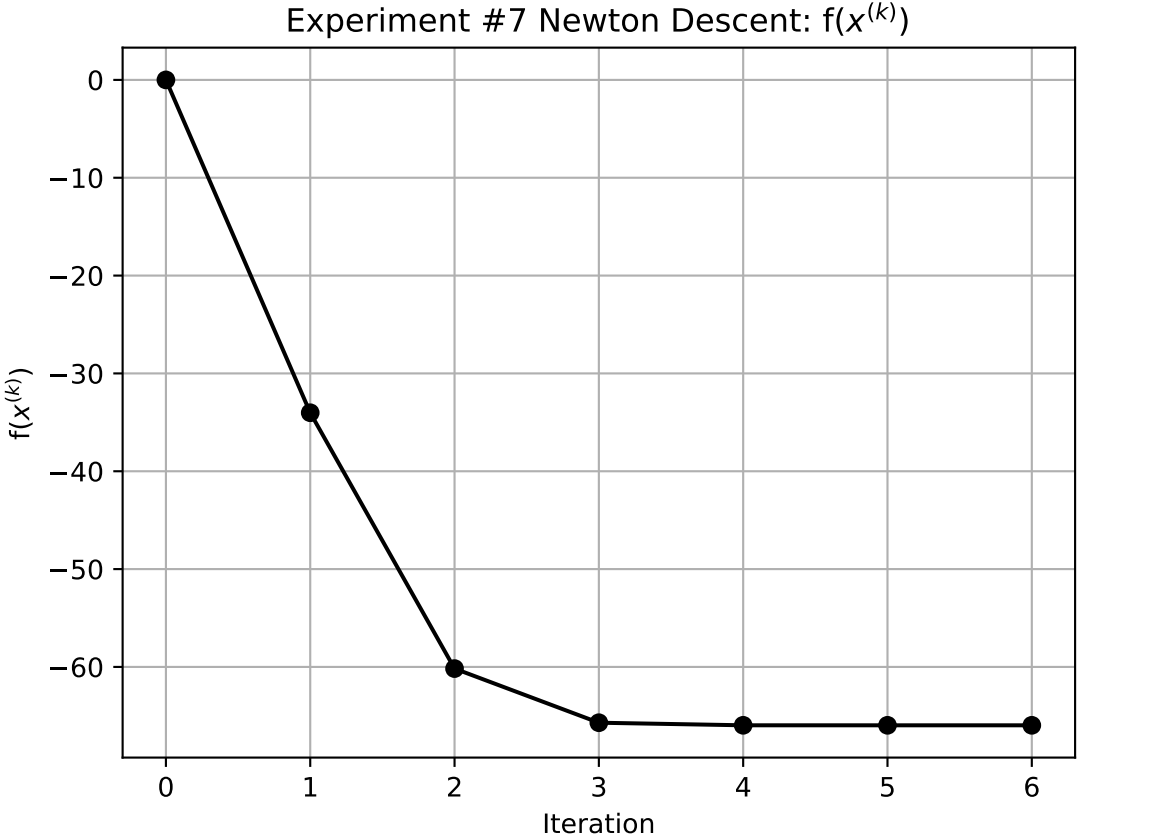

Experiment #5 Newton Descent: Error f($x^{(k)}$) - p^*
f($x^{(k)}$) - p^*
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
0
2
4
6
8
10
12
Iteration

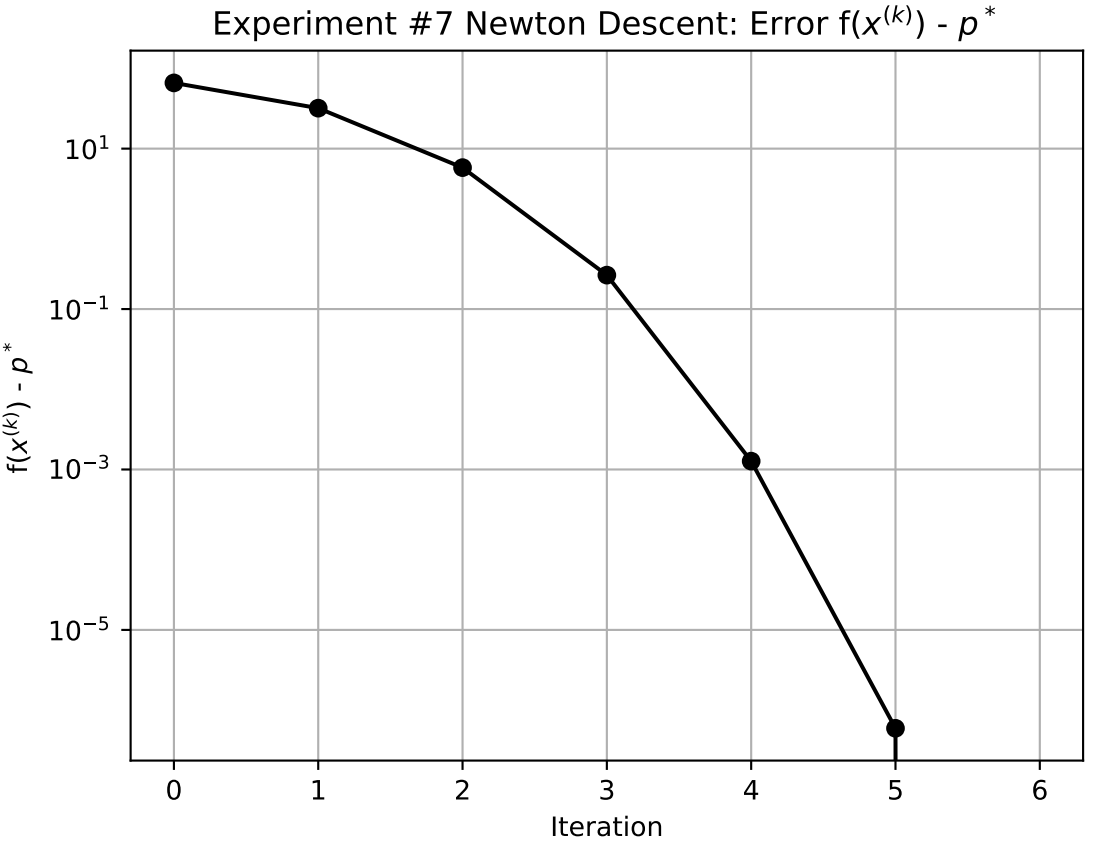

Experiment #5 Newton Descent: Step Size
10^0
6×10^{-1}
4×10^{-1}
3×10^{-1}
Step Size
0
2
4
6
8
10
12
Iteration

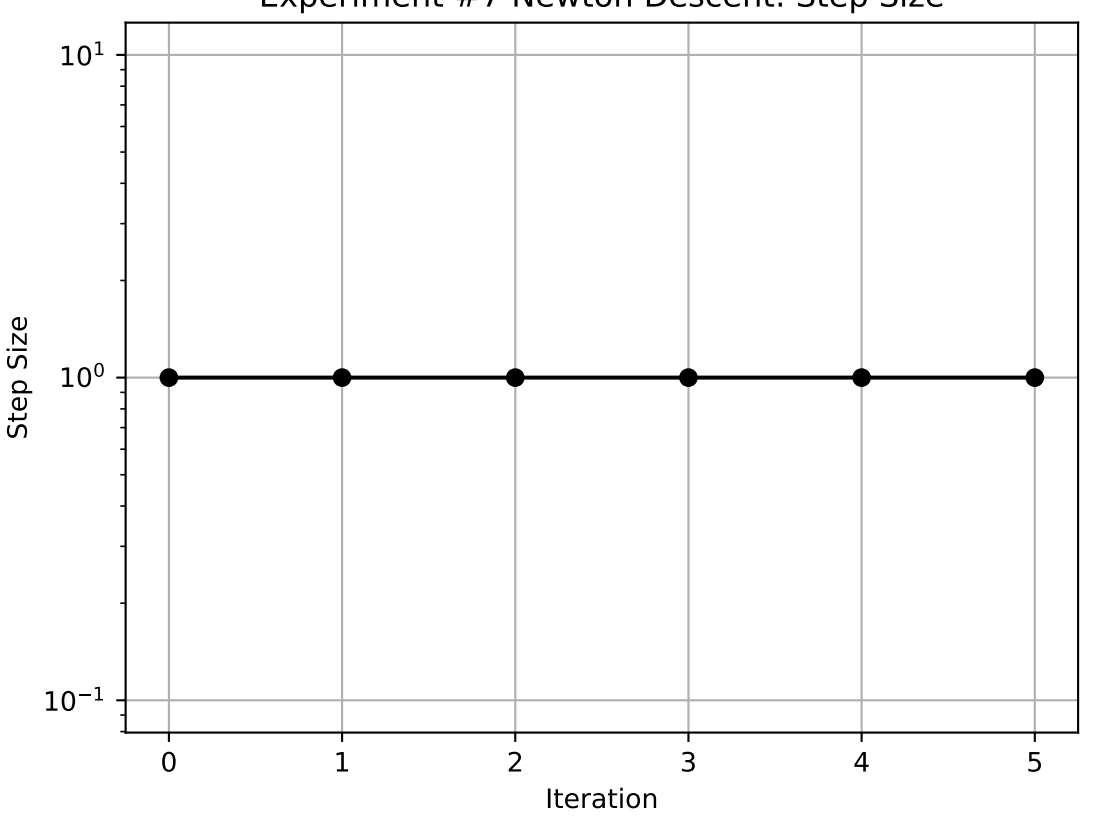

Experiment #6 Newton Descent: f($x^{(k)}$)
0
−50
−100
−150
−200
−250
−300
−350
−400
f($x^{(k)}$)
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Iteration

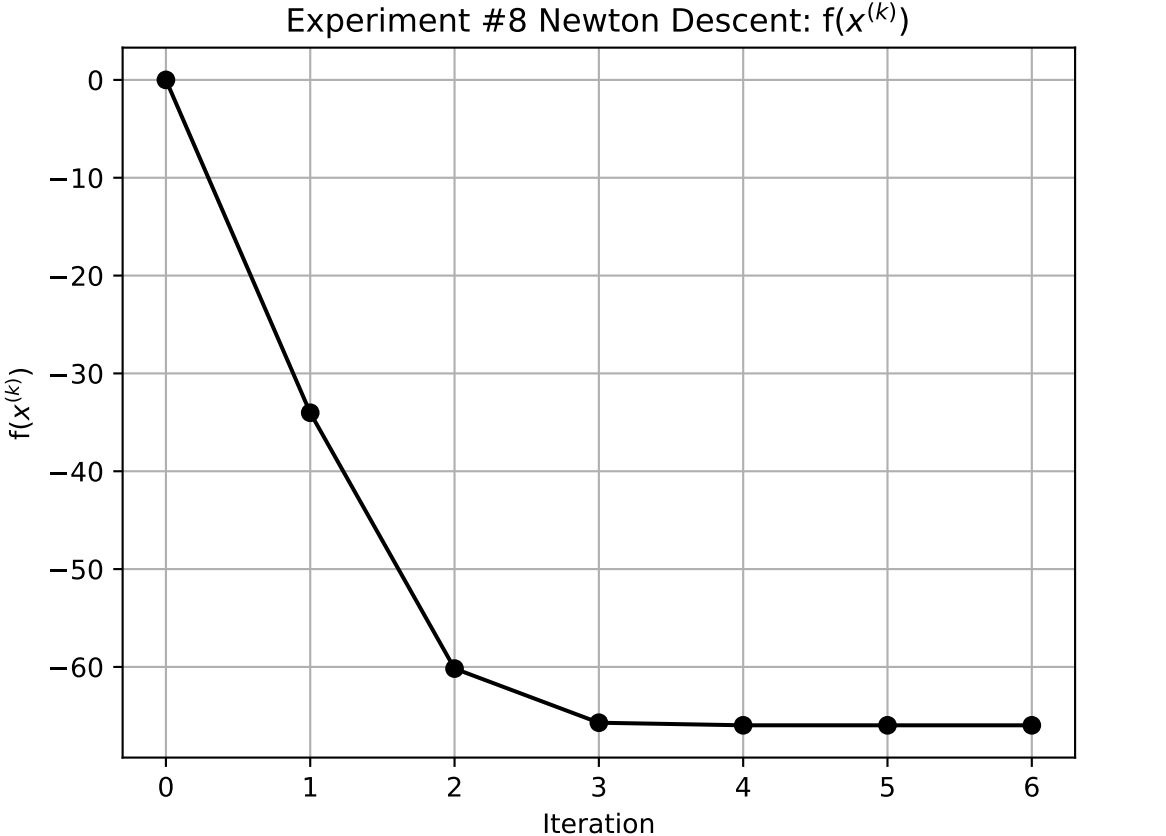

Experiment #6 Newton Descent: Error f($x^{(k)}$) - p^*
f($x^{(k)}$) - p^*
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Iteration

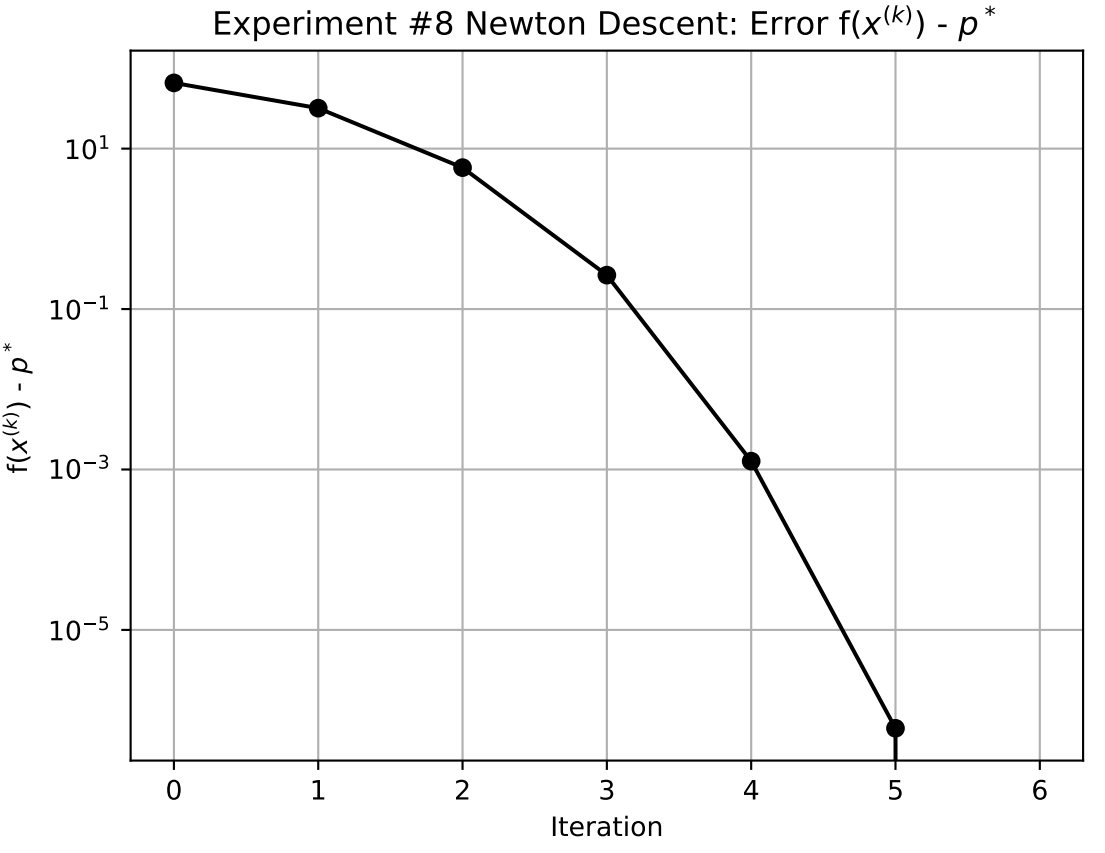

Experiment #6 Newton Descent: Step Size
10^0
10^{-1}
Step Size
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Iteration

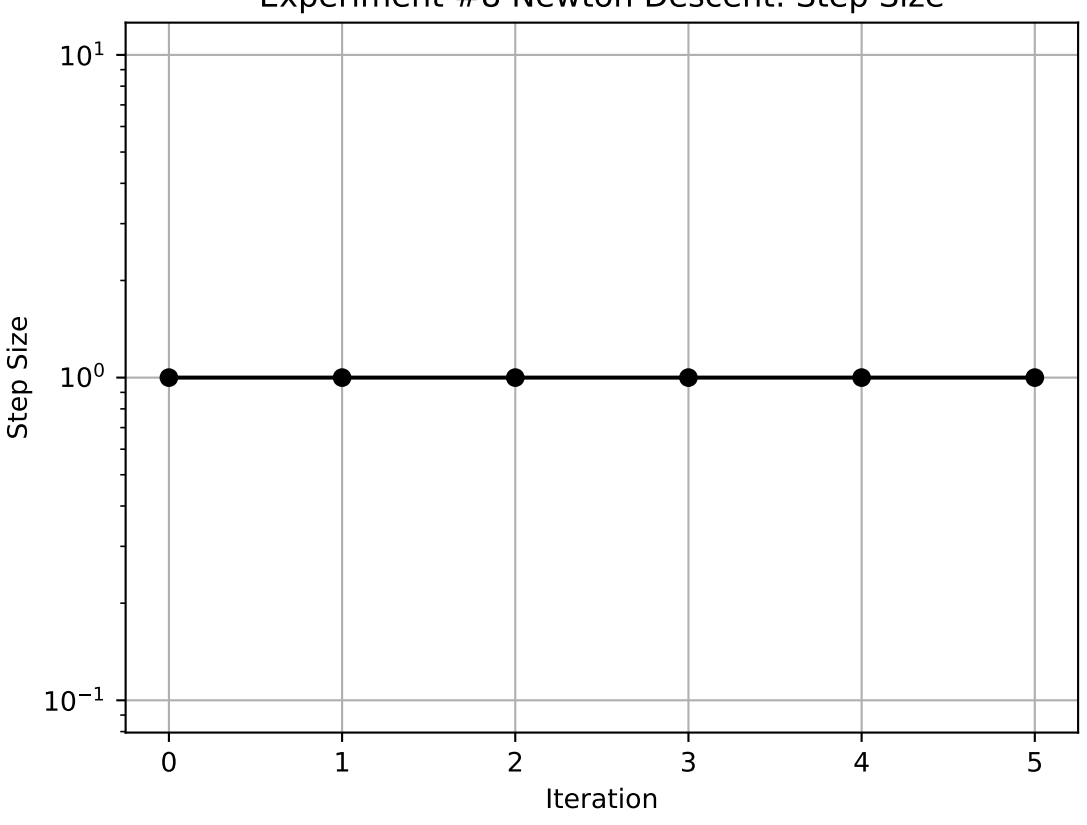

Experiment #7 Newton Descent: f($x^{(k)}$)
0
−10
−20
−30
−40
−50
−60
f($x^{(k)}$)
0
1
2
3
4
5
6
Iteration

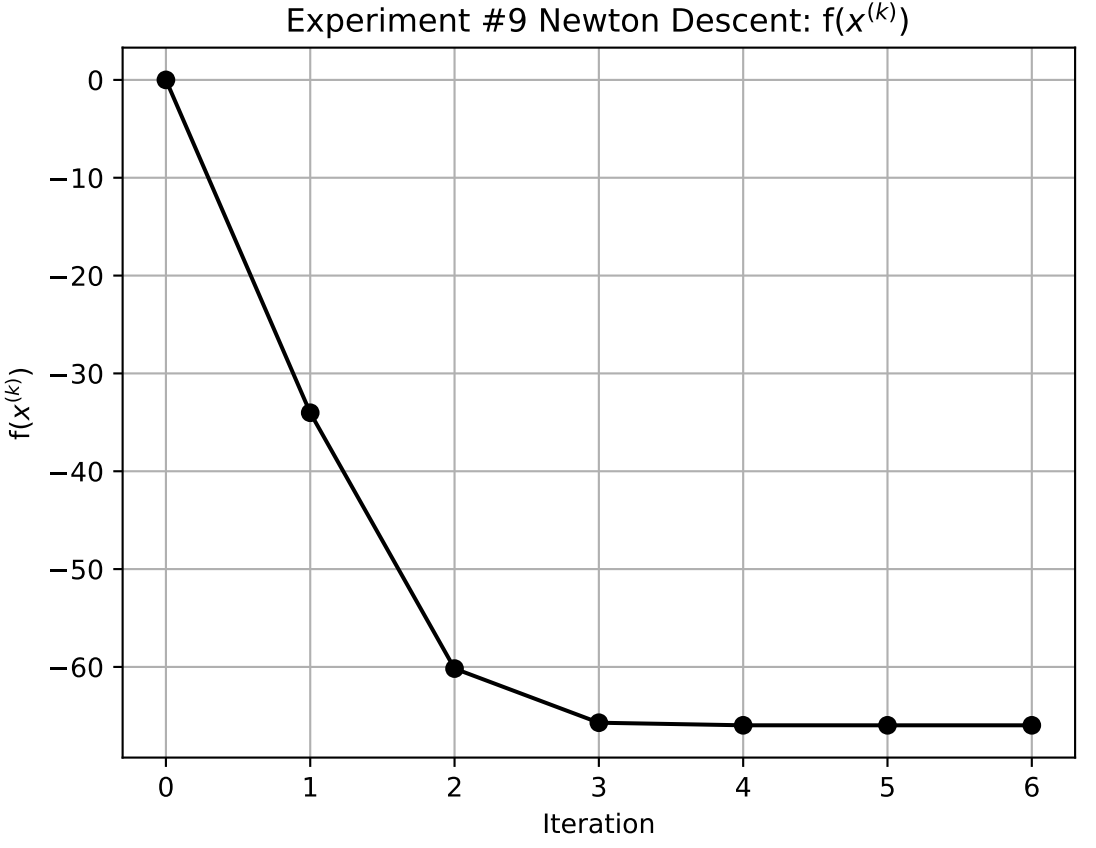

Experiment #7 Newton Descent: Error f($x^{(k)}$) - p^*
10^1
10^{-1}
10^{-3}
10^{-5}
f($x^{(k)}$) - p^*
0
1
2
3
4
5
6
Iteration

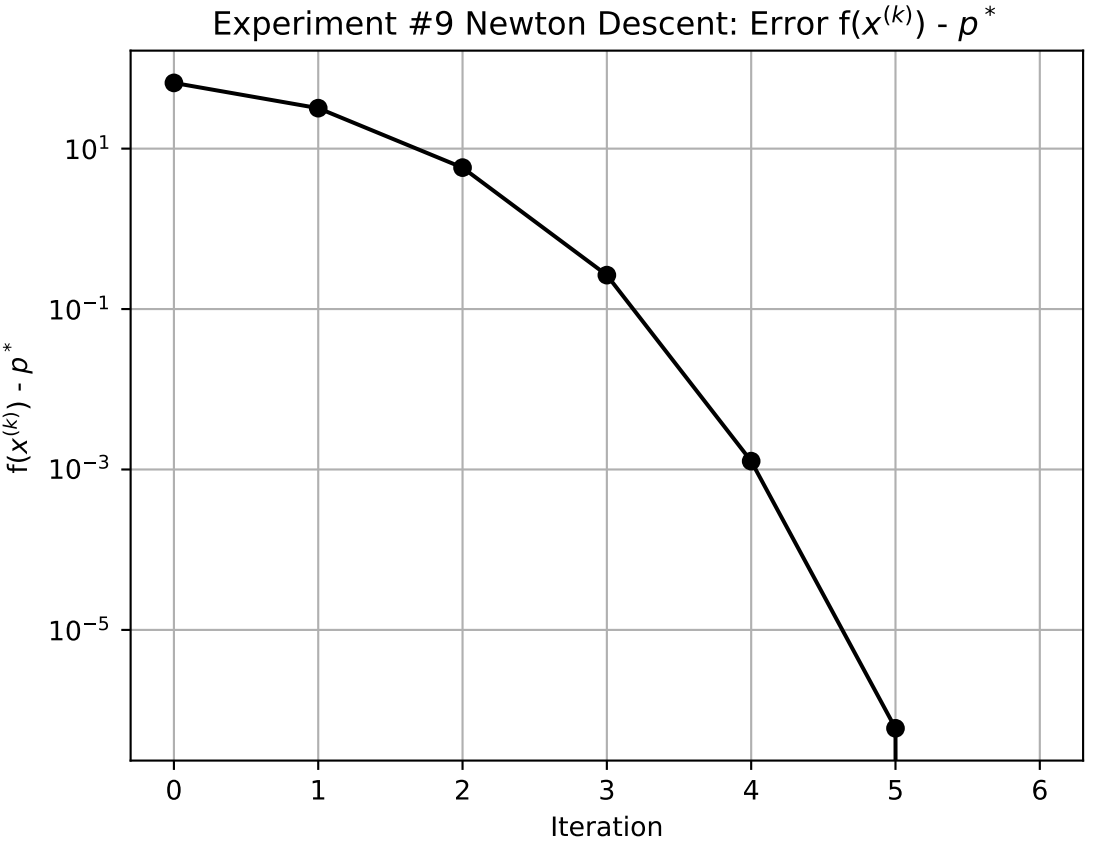

Experiment #7 Newton Descent: Step Size
Step Size
10^{1}
10^{0}
10^{-1}
0
1
2
3
4
5
Iteration

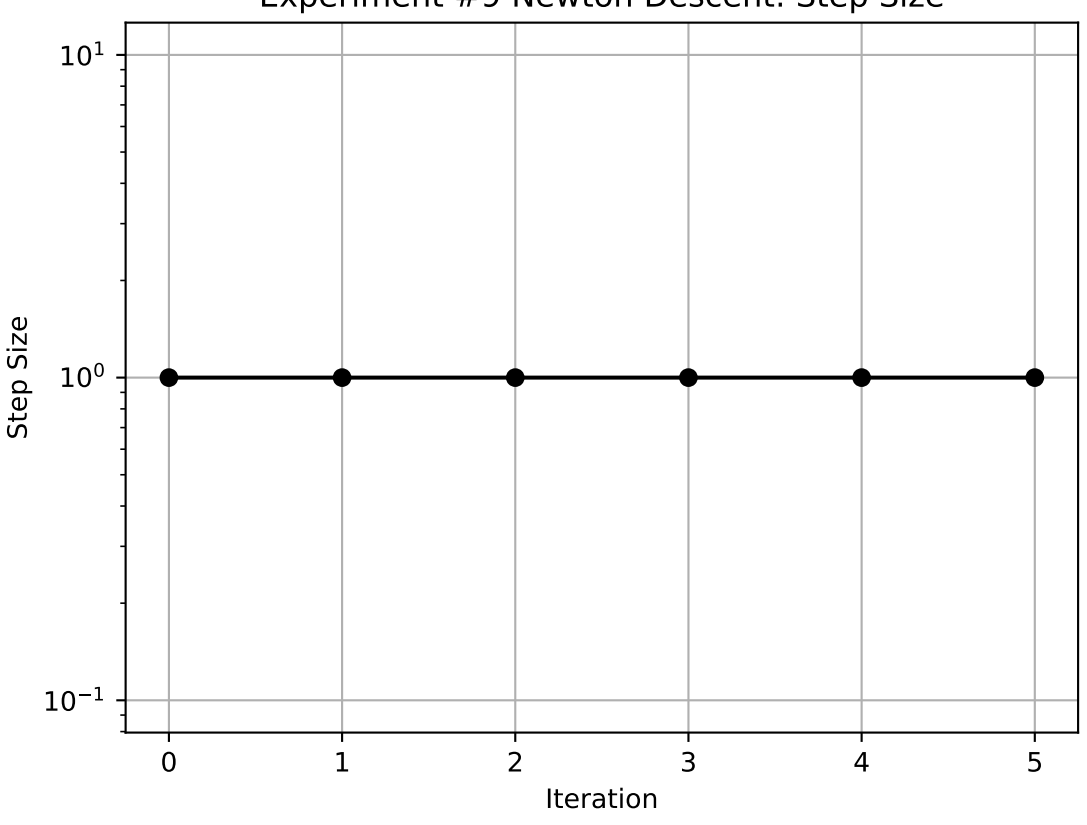

Experiment #8 Newton Descent: f($x^{(k)}$)
0
−10
−20
−30
−40
−50
−60
f($x^{(k)}$)
0
1
2
3
4
5
6
Iteration

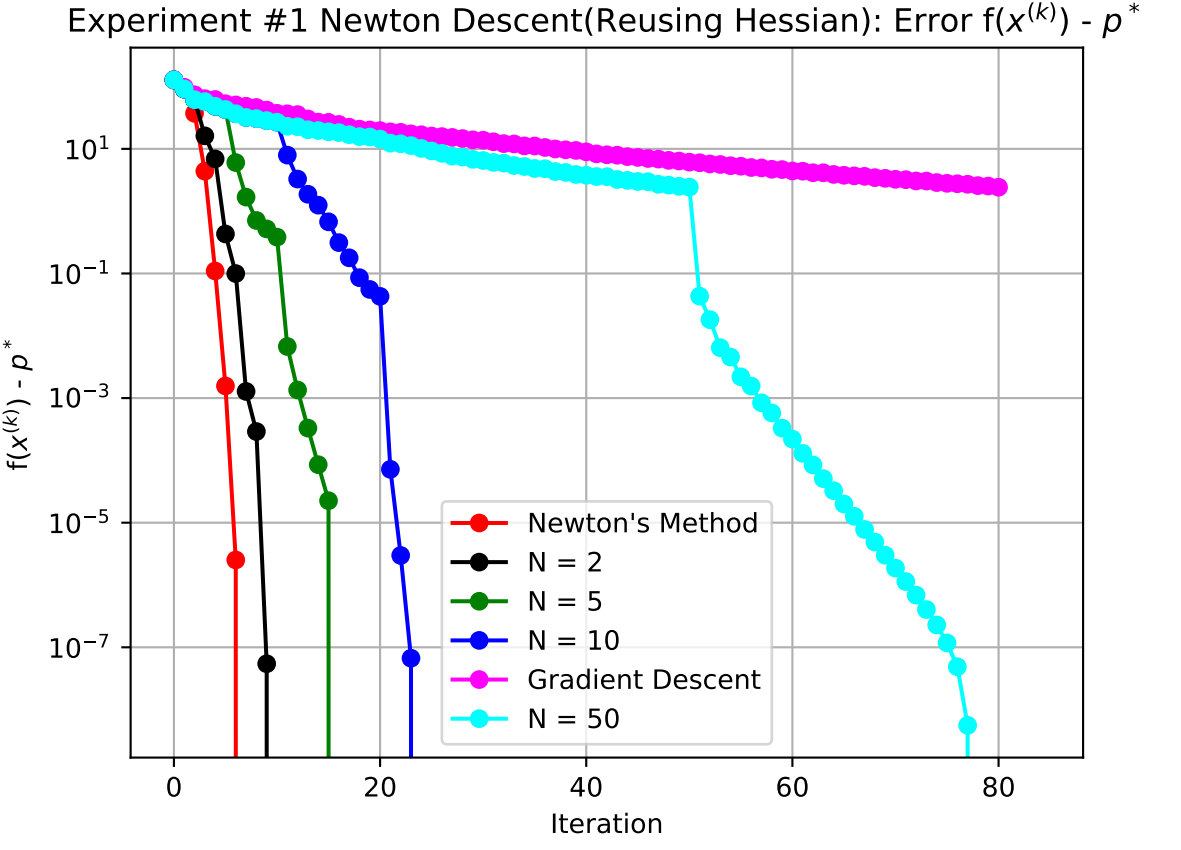

Experiment #8 Newton Descent: Error f($x^{(k)}$) - p^*
10^1
10^{-1}
10^{-3}
10^{-5}
f($x^{(k)}$) - p^*
0
1
2
3
4
5
6
Iteration

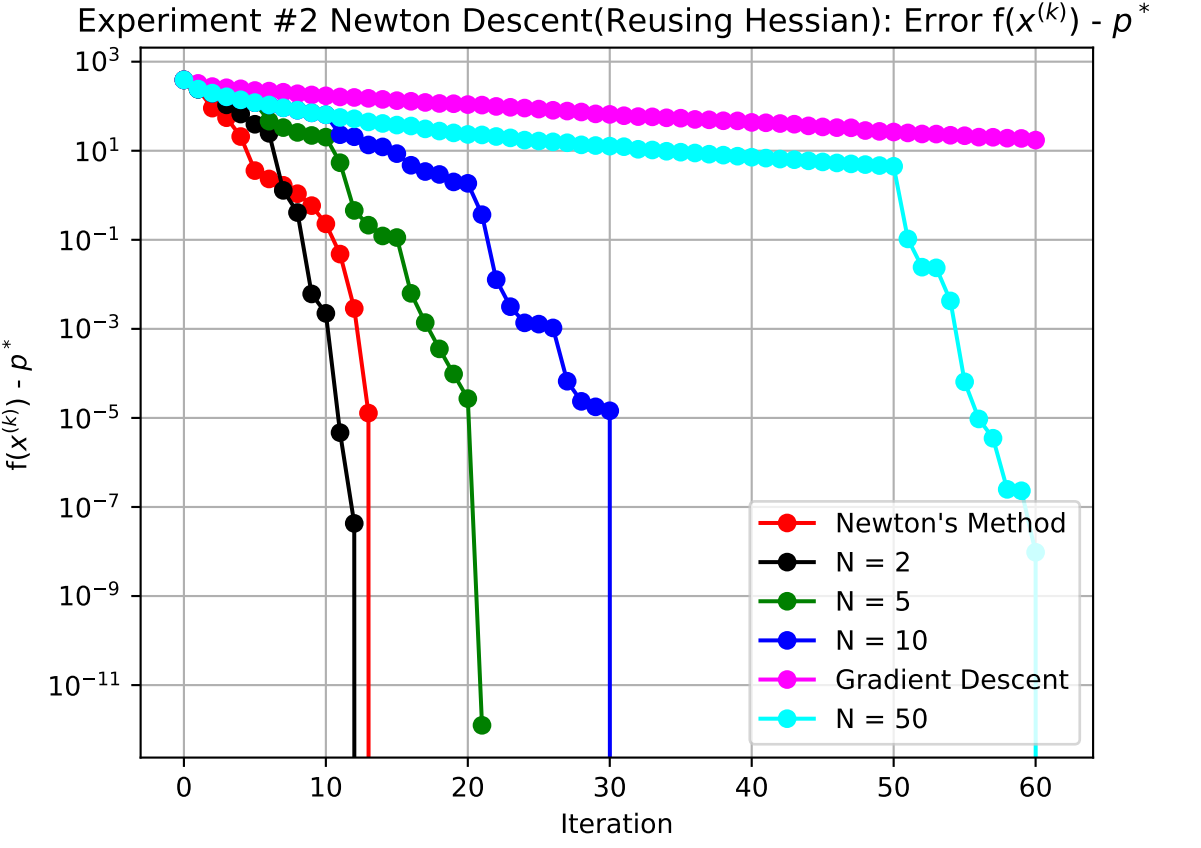

Experiment #8 Newton Descent: Step Size
Step Size
10^{1}
10^{0}
10^{-1}
0
1
2
3
4
5
Iteration

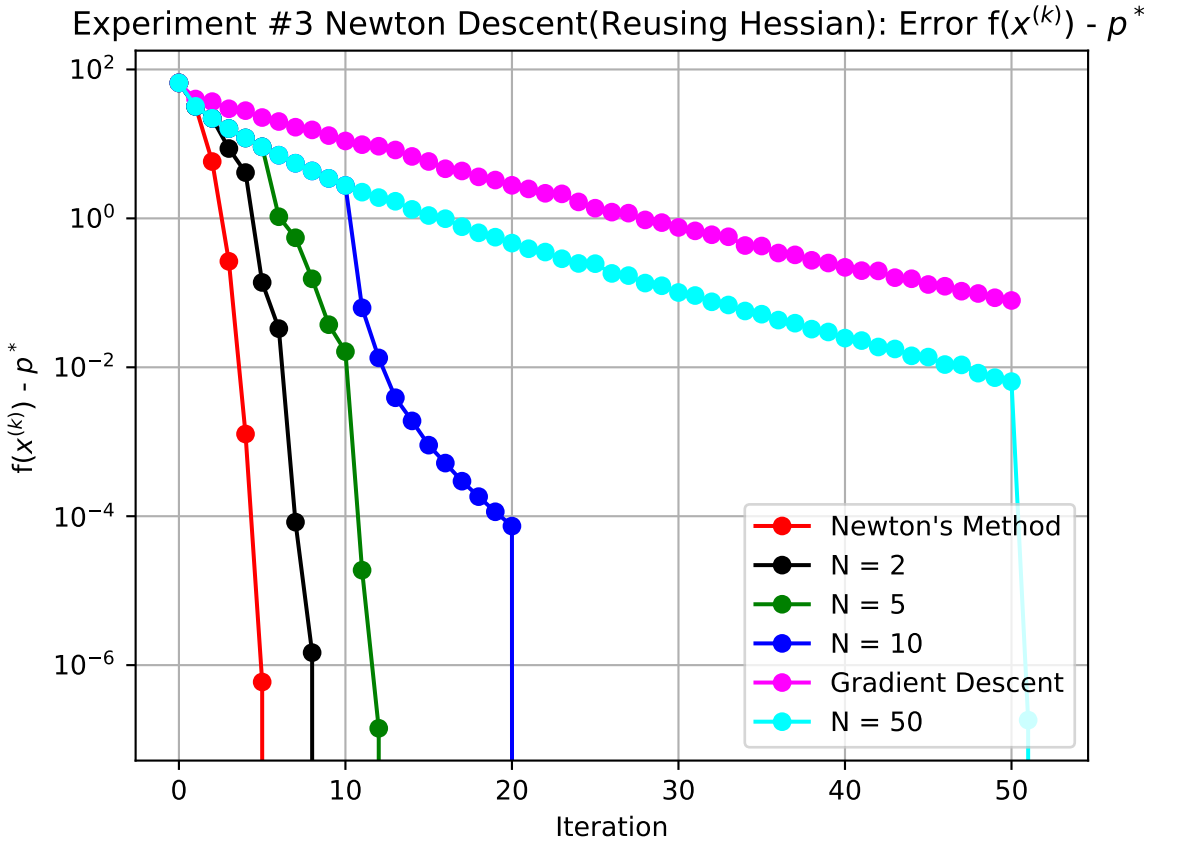

Experiment #9 Newton Descent: f($x^{(k)}$)
0
−10
−20
−30
−40
−50
−60
f($x^{(k)}$)
0
1
2
3
4
5
6
Iteration

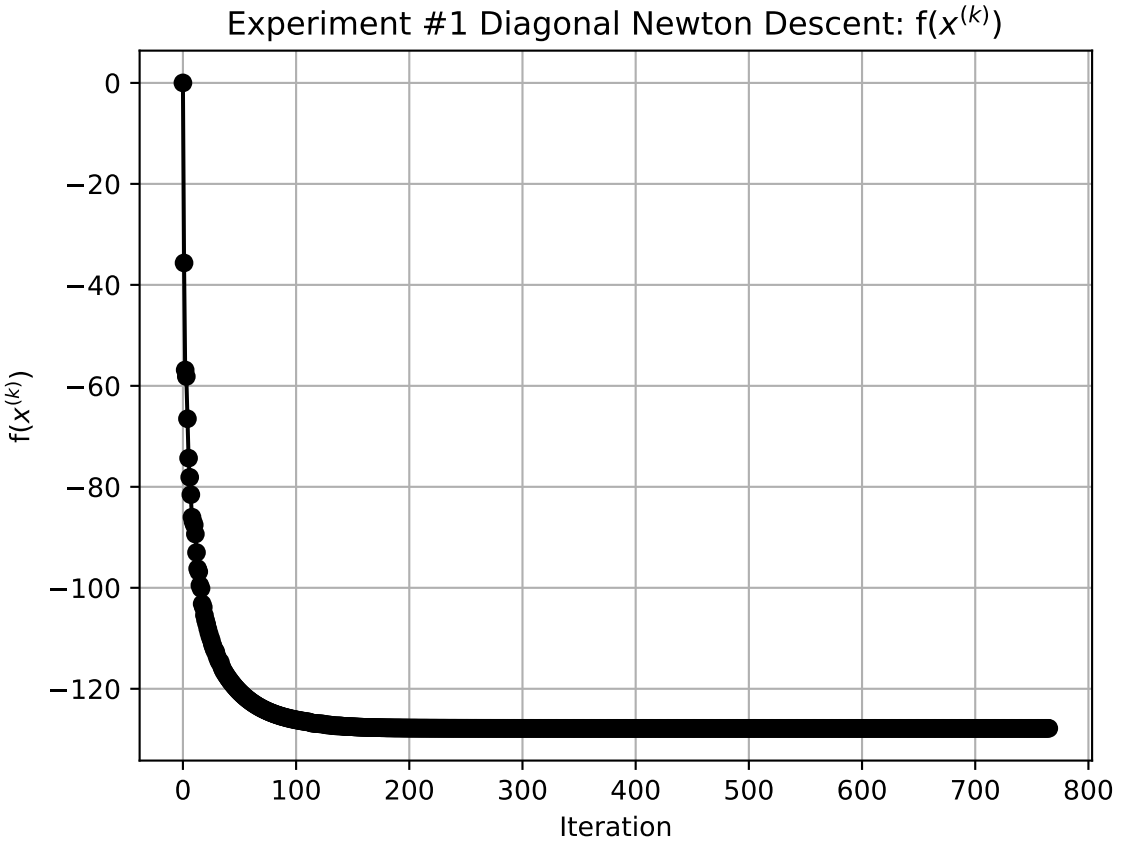

Experiment #9 Newton Descent: Error f($x^{(k)}$) - p^*
10^1
10^{-1}
10^{-3}
10^{-5}
f($x^{(k)}$) - p^*
0
1
2
3
4
5
6
Iteration

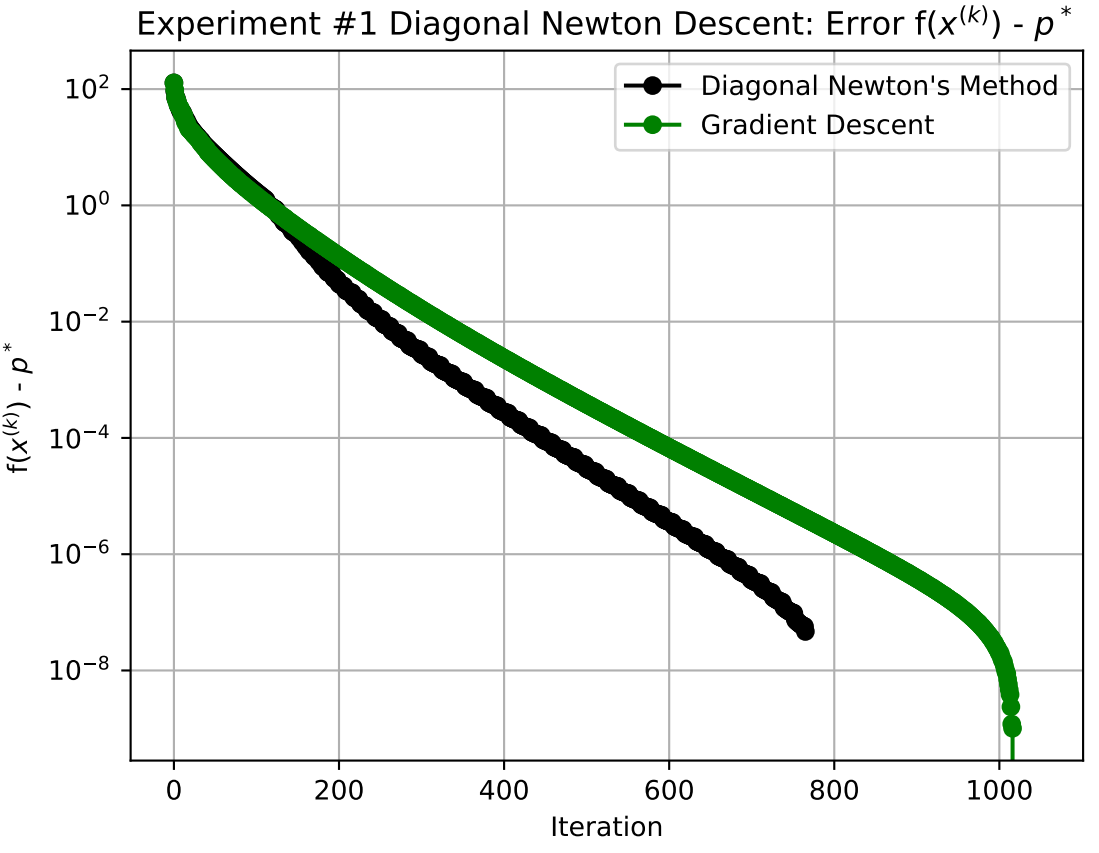

Experiment #9 Newton Descent: Step Size
Step Size
10^{1}
10^{0}
10^{-1}
0
1
2
3
4
5
Iteration

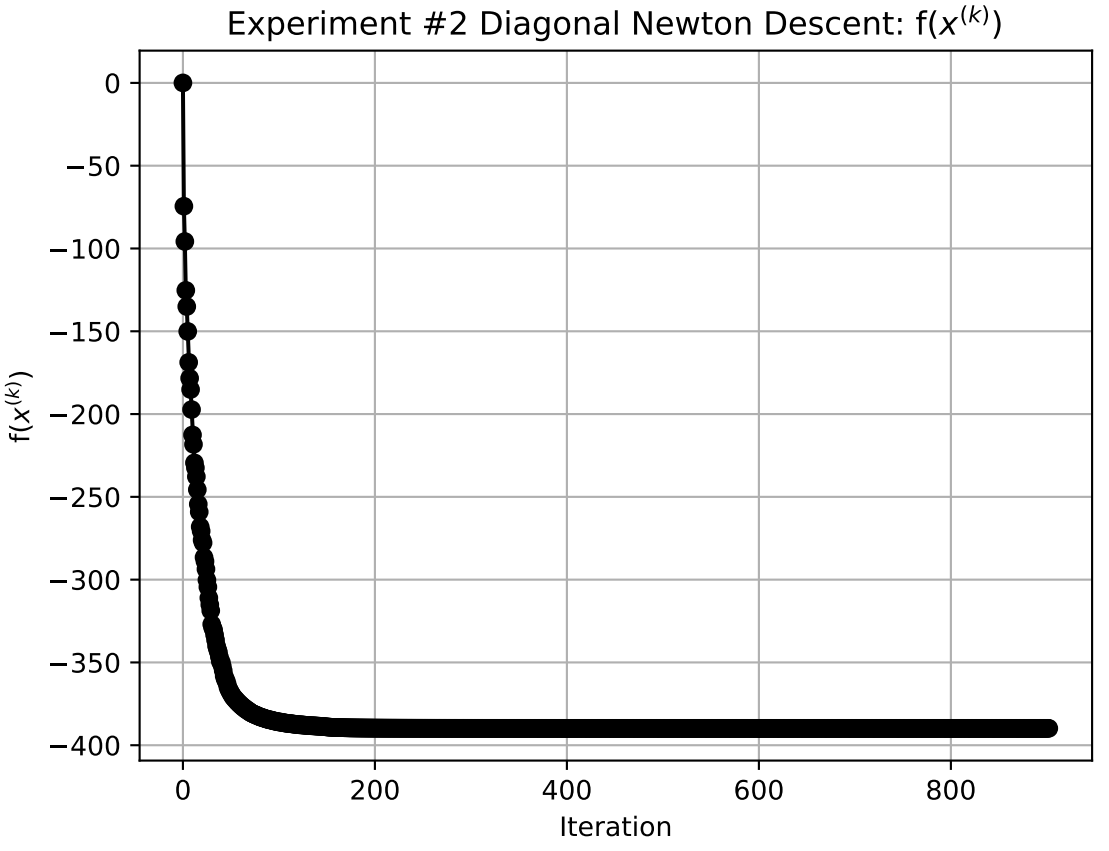

Experiment #1 Newton Descent(Reusing Hessian): Error f($x^{(k)}$) - p^*
f($x^{(k)}$) - p^*
10^{1}
10^{-1}
10^{-3}
10^{-5}
10^{-7}
0
20
40
60
80
Iteration
Newton's Method
N = 2
N = 5
N = 10
Gradient Descent
N = 50

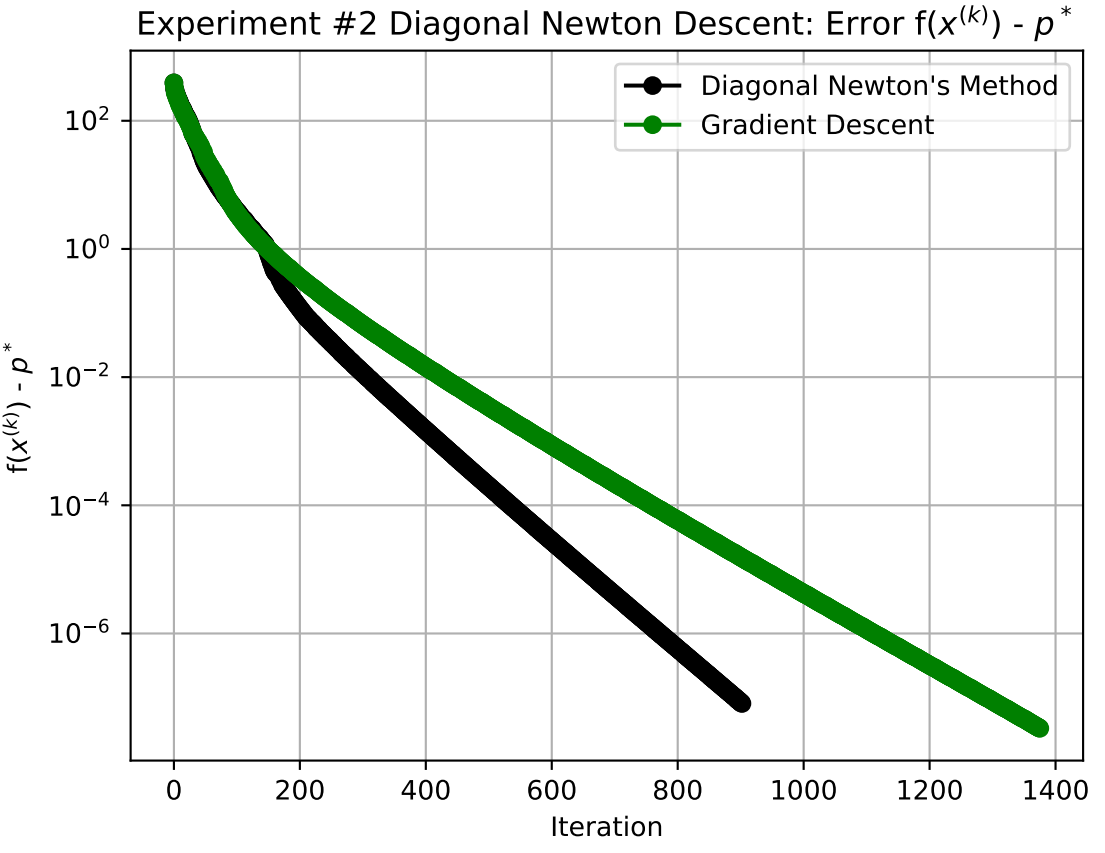

Experiment #2 Newton Descent(Reusing Hessian): Error f($x^{(k)}$) - p^*
10^3
10^1
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-9}
10^{-11}
f($x^{(k)}$) - p^*
0
10
20
30
40
50
60
Iteration
Newton's Method
N = 2
N = 5
N = 10
Gradient Descent
N = 50

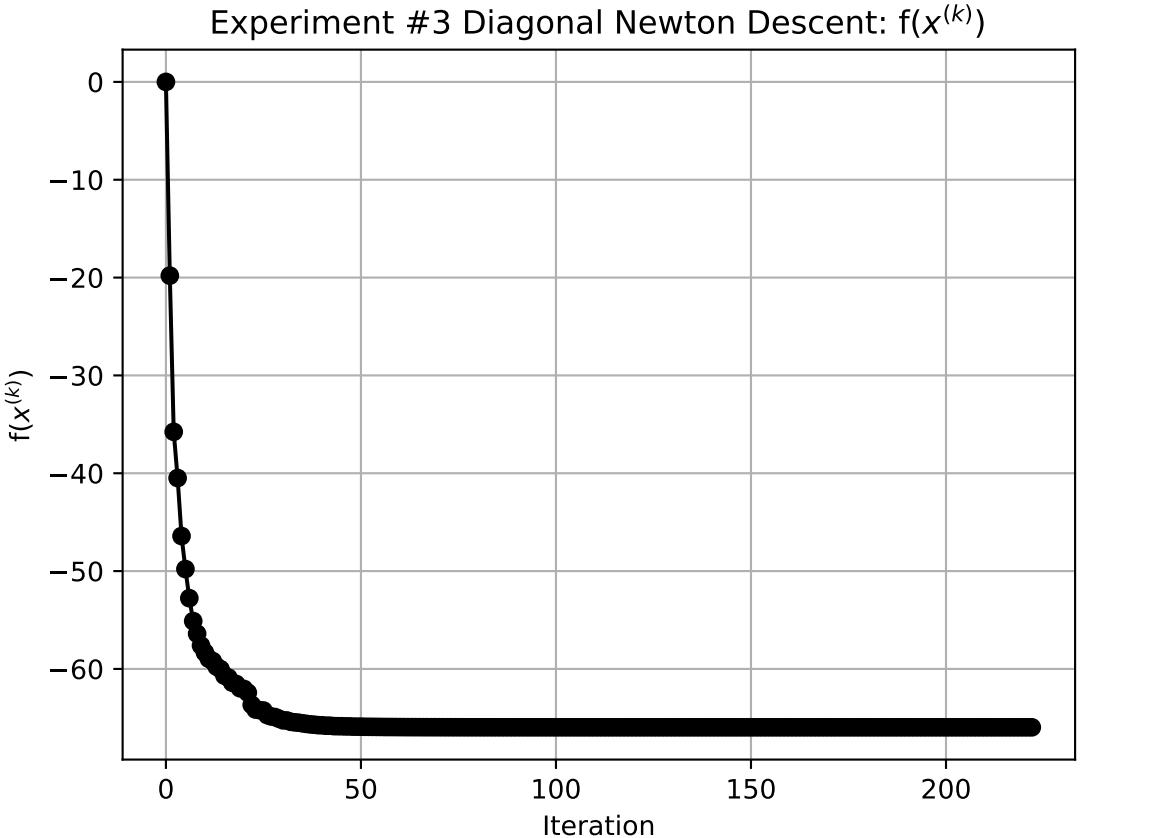

Experiment #3 Newton Descent(Reusing Hessian): Error f($x^{(k)}$) - p^*
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
f($x^{(k)}$) - p^*
0
10
20
30
40
50
Iteration
Newton's Method
N = 2
N = 5
N = 10
Gradient Descent
N = 50

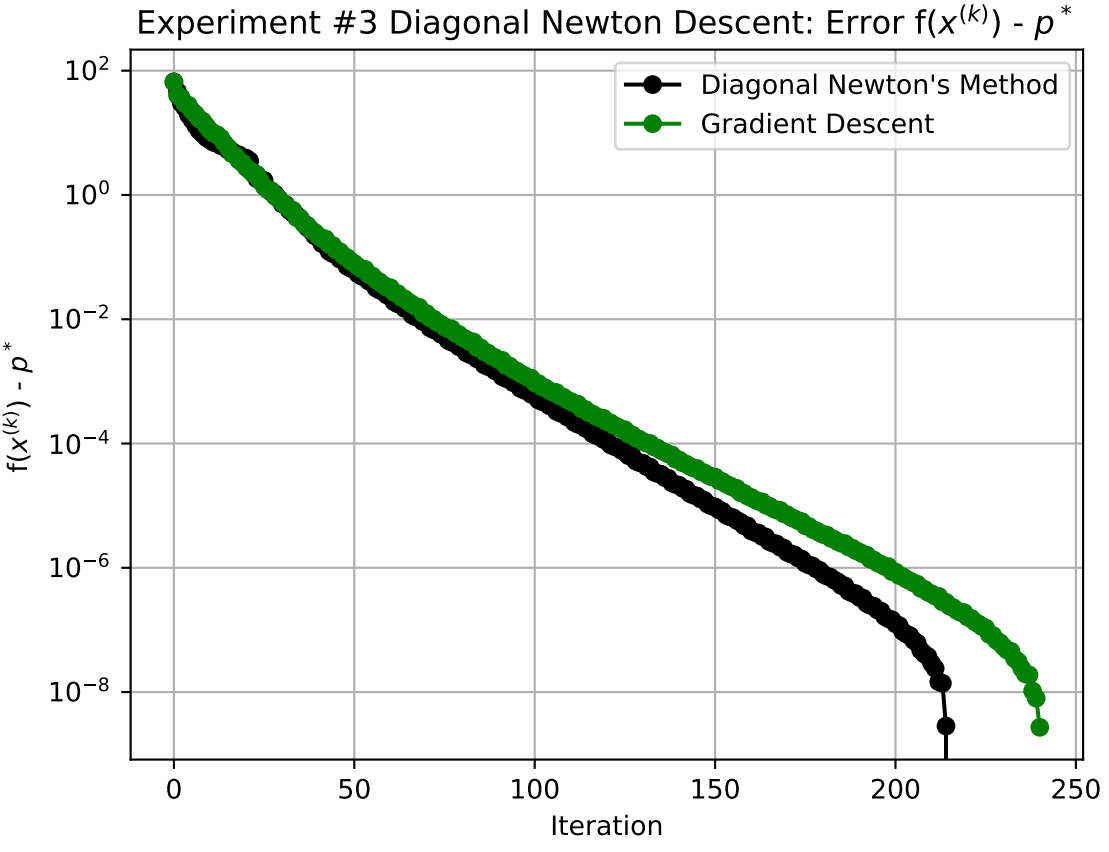

Experiment #1 Diagonal Newton Descent: f($x^{(k)}$)
0
−20
−40
−60
−80
−100
−120
f($x^{(k)}$)
0
100
200
300
400
500
600
700
800
Iteration


Experiment #1 Diagonal Newton Descent: Error f($x^{(k)}$) - p^*
Diagonal Newton's Method
Gradient Descent
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
f($x^{(k)}$) - p^*
0
200
400
600
800
1000
Iteration


Experiment #2 Diagonal Newton Descent: f($x^{(k)}$)
0
−50
−100
−150
−200
−250
−300
−350
−400
f($x^{(k)}$)
0
200
400
600
800
Iteration


Experiment #2 Diagonal Newton Descent: Error f($x^{(k)}$) - p^*
Diagonal Newton's Method
Gradient Descent
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
f($x^{(k)}$) - p^*
0
200
400
600
800
1000
1200
1400
Iteration


Experiment #3 Diagonal Newton Descent: f($x^{(k)}$)
0
−10
−20
−30
−40
−50
−60
f($x^{(k)}$)
0
50
100
150
200
Iteration


Experiment #3 Diagonal Newton Descent: Error f($x^{(k)}$) - p^*
Diagonal Newton's Method
Gradient Descent
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
f($x^{(k)}$) - p^*
0
50
100
150
200
250
Iteration